

AVERAGED DYNAMICS OF ULTRA-RELATIVISTIC CHARGED PARTICLE BEAMS

by

Ricardo Gallego Torromé

This thesis submitted for the degree of Doctor of Philosophy

Abstract

In this thesis, we consider the suitability of using the charged cold fluid model in the description of ultra-relativistic beams. The method that we have used is the following. Firstly, the necessary notions of kinetic theory and differential geometry of second order differential equations are explained. Then an averaging procedure is applied to a connection associated with the Lorentz force equation. The result of this averaging is an affine connection on the space-time manifold. The corresponding geodesic equation defines the averaged Lorentz force equation. We prove that for ultra-relativistic beams described by narrow distribution functions, the solutions of both equations are similar. This fact justifies the replacement of the Lorentz force equation by the simpler *averaged Lorentz force equation*. After this, for each of these models we associate the corresponding kinetic model, which are based on the Vlasov equation and *averaged Vlasov equation* respectively. The averaged Vlasov equation is simpler than the original Vlasov equation. This fact allows us to prove that the differential operation defining the averaged charged cold fluid equation is controlled by the *diameter of the distribution function*, by powers of the *energy of the beam* and by the time of evolution t . We show that the Vlasov equation and the averaged Vlasov equation have similar solutions, when the initial conditions are the same. Finally, as an application of the *averaged Lorentz force equation* we re-derive the beam dynamics formalism used in accelerator physics from the Jacobi equation of the averaged Lorentz force equation.

Acknowledgements.

I would like to express my gratitude to Dr. Volker Perlick for the many discussions on the material of this thesis, his advice and patience in the supervision of the thesis. I would also like to acknowledge to the members of the mathematical physics group of the physics department at Lancaster University for many discussions and advice in some matters concerning the material presented in this memoir and to the Cockcroft Institute for support. My friends at Lancaster have also contributed to make this thesis possible and I express to them my gratitude.

This thesis is dedicated to my parents, Luis and Elvira and to my brothers Luis Miguel, Julio and Elvira.

Contents

1	Introduction	10
1.1	Motivation of the thesis	10
1.1.1	Other results of the thesis	13
1.2	Structure of the thesis	14
1.3	General conventions used in the thesis	16
2	Fluid and kinetic models for ultra-relativistic beams	20
2.1	Basic relativistic kinetic theory	20
2.1.1	Intrinsic covariant formalism for relativistic kinetic theory	20
2.1.2	Extrinsic formulation of the kinetic model	23
2.2	Relativistic charged cold fluid model	25
2.3	Asymptotic expansion of the relativistic charged cold fluid model . .	27
2.4	Statement of the main problem considered in this thesis and out-line of the strategy to solve it	29
3	The averaged connection	31
3.1	Non-Linear connection associated with a second order differential equation	31
3.1.1	Second order differential equations and the associated non- linear Berwald-type connection	31

3.1.2	The pull-back bundle	34
3.2	The average operator associated with a family of automorphisms . .	36
3.2.1	Average of a family of automorphisms	36
3.2.2	Examples of geometric structures which provide an averaging procedure	38
3.2.3	Average operator acting on sections	42
3.3	Averaged connection of a linear connection on $\pi^*\mathbf{TM}$	43
4	Comparison of the Lorentz force equation and the averaged Lorentz force equation	47
4.1	Introduction	47
4.1.1	On the physical interpretation of the formalism	50
4.2	Criticism of the notion of semi-Randers space as space-time structure	52
4.2.1	Randers spaces as space-time structures	52
4.2.2	Comparison of Asanov's and Beem's formalism	53
4.3	Non-Lagrangian notion of semi-Randers space	56
4.3.1	Non-Lagrangian notion of semi-Randers space	56
4.3.2	Variational principle on semi-Randers spaces	58
4.4	Geometric formulation of the Lorentz force equation	63
4.4.1	The non-linear connection associated with the Lorentz force equation	63
4.4.2	The Koszul connection ${}^L\nabla$ on \mathbf{TN} associated with the Lorentz force equation	65
4.4.3	The Lorentz connection on the pull-back bundle $\pi^*\mathbf{TM}$. . .	69
4.5	The averaged Lorentz connection	71
4.6	Comparison between the geodesics of ${}^L\nabla$ and $\langle {}^L\nabla \rangle$	74
4.6.1	Basic geometry in the space of connections	74

4.6.2	Comparison between the geodesics of ${}^L\nabla$ and $\pi^* < {}^L\nabla >$. . .	82
4.7	Discussion	88
4.7.1	Structural stability of the approximation ${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$	88
4.7.2	On the hypotheses on which the approximation ${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$ is based	91
4.7.3	Range of applicability of the approximation	93
5	Charged cold fluid model from the Vlasov model	96
5.1	Introduction	96
5.2	Comparison of the solutions of the Vlasov and averaged Vlasov equations	99
5.2.1	Examples of Liouville equations	99
5.2.2	Comparison between the solutions of the Vlasov equation and the averaged Vlasov equation	101
5.3	The charged cold fluid model from the averaged Vlasov model	104
5.3.1	Comparison of the Vlasov model with the averaged Vlasov model	105
5.3.2	Bound on the auto-parallel condition of the unitary mean vector field of the averaged Vlasov model	114
5.3.3	Bound on the auto-parallel condition of the mean velocity field of the Vlasov model	116
5.4	Discussion	117
5.4.1	On the validity of the truncation schemes in fluid models . . .	118
6	The Jacobi equation of the averaged Lorentz connection and applications in beam dynamics	121
6.1	Introduction	121
6.1.1	The Jacobi equation of an affine connection on \mathbf{M}	121

6.1.2	Jacobi equation for linear connections defined on the pull-back bundle $\pi^*\mathbf{TM} \rightarrow \mathbf{N}$	123
6.1.3	Physical Interpretation of the Jacobi equation	125
6.2	Jacobi equation of the averaged lorentz connection	126
6.2.1	The Jacobi equation of the Lorentz connection versus the Jacobi equation of the averaged Lorentz connection	128
6.3	Transversal beam dynamics from the Jacobi equation of the averaged connection	129
6.3.1	Relation between the transverse dynamics and the Jacobi equation	129
6.3.2	Examples of transverse linear dynamics	132
6.4	Calculation of the averaged off-set effect between the reference trajectory and the central geodesic of the averaged Lorentz connection .	135
6.4.1	Calculation of the dispersion function in beam dynamics . . .	135
6.4.2	Calculation of the off-set due to the deviation ϵ^j	137
6.5	Longitudinal beam dynamics and corrections from the Jacobi equation of the averaged connection	141
6.5.1	Examples	143
6.6	Conclusion	144
7	Conclusions	146
7.1	General conclusions	146
7.1.1	Brief discussion of the main problem	148
7.2	Generalizations and open problems	149
A	Mathematical appendix	151
A.1	Proofs for Chapter 3	151
A.2	Ordinary differential equations	155
A.3	Basic notions of asymptotic analysis	156

A.4	Notion of almost-connection	158
A.5	Basic notions of Sobolev spaces	162
A.5.1	Basic properties of Sobolev spaces	163
A.5.2	Sobolev spaces of functions defined on manifolds	164
A.5.3	Examples of Sobolev spaces	164

Chapter 1

Introduction

1.1 Motivation of the thesis

Current models of classical electrodynamics of charged point particles contain logical inconsistencies that arise when back-reaction effects are considered. For example, the standard theory of back-reaction is based on the Lorentz-Dirac equation [1-5]. However, it is well known that the Lorentz-Dirac equation is problematic from a physical point of view: some of its solutions contain pre-acceleration effects; others are run-away solutions. This is the case for a large class of initial conditions. This peculiarity of the Lorentz-Dirac equation is due to the fact that it is a third order differential equation.

A possible solution to the problems of the Lorentz-Dirac equation is the theory proposed by Landau and Lifshitz in [2] (recently reviewed for instance in reference [5]). From the analysis of this question performed in reference [5], one extracts the following conclusion:

The charged point particle description is valid iff the changes in the acceleration of the particle occur over time scales longer than the characteristic time parameter $t_0 = \frac{2}{3} \frac{q^2}{m}$.

The parameter q is the charge and m is the mass of the point particle and one has assumed units such that the permeability of the vacuum ϵ_0 is 1 and the speed of

light c is also set equal to 1. If the point particle approximation condition holds, the Landau-Lifshitz reduction of order procedure can be applied to the Lorentz-Dirac equation to obtain a second order differential equation as an approximation, free from pathological solutions. In this regime, the Landau-Lifshitz equation can be considered an appropriate approximation of the Lorentz-Dirac equation.

Despite solving the problem of the Lorentz-Dirac equation, there are several reasons why the solution proposed by Landau and Lifshitz is not completely satisfactory:

1. The *order reduction procedure* is an ad-hoc procedure (although consistent with the point particle picture).
2. The Landau-Lifshitz equation is the leading order term approximation of the Lorentz-Dirac equation. Therefore, it is not a fundamental equation.
3. The characteristic time t_0 is proportional to $\frac{q}{m}q$. Therefore, let us consider a physical system with a large number of identical charged particles performing a *collective motion*. The prototype example is the motion of a bunch of particles in an accelerator machine. It can happen that the behavior of the system is coherent and that one has to read the factor q as the total charge of the bunch and m as the total mass of the whole bunch. Under these conditions, it is natural to consider that the factor $\frac{q}{m}$ remains the same as for an individual charged point particle, but q increases proportionally with the number of particles. Then for intense beams of particles, the point charge approximation and the reduction of order procedure will break down.

Even if the first two points can be *covered* under the interpretation of classical electrodynamics as the limit of the fundamental quantum electrodynamics, the third point has relevance for us. The energies and luminosity achieved in modern particle accelerators can push to the validity of present models of electrodynamics its limits. This is basically because one is dealing with bunches containing a large number of charged particles, which can reach $10^9 - 10^{11}$ particles per bunch, moving together in a small phase-space domain (all the particles are concentrated around a center of

mass, in position and velocity).

Since the possible effect discussed in point 3 is additive, for modeling systems like those bunches of particles, one needs an alternative description to Lorentz-Dirac and Landau-Lifshitz models.

In this context, fluid models have been used to study the dynamics of ultra-relativistic beams of charged particles. One of these models is the proposal contained in [6]. In that work, it was shown how to do an asymptotic analysis of the charged cold fluid model. The main claim in [6] was that the model proposed provides a self-consistent description of the fully coupled dynamics of a bunch of particles with the electromagnetic field. The reason for this is the smoothness properties of the fields, compared with the discrete and singular character of the point particle description behind the Lorentz-Dirac equation.

However, the use of the charged cold fluid model was not justified in [6]. This justification is necessary, because of the discrete nature of a bunch of particles. Therefore, prior to the use of this model, one has to address the following question:

When is it a good approximation, in the regime of ultra-relativistic dynamics, to describe the interaction of a large number of charged point particles with the total (external and associated) electromagnetic field by a charged cold fluid model?

A simplified version, is the following question:

When is it a good approximation, in the regime of ultra-relativistic dynamics, to describe the interaction of a large number of charged point particles with an external electromagnetic field by a charged cold fluid model?

In the present thesis we address this second question. In particular, we present an *averaged description* of the collection of charged point particles, defining a *mean velocity vector field*. The averaging operation is interpreted from a kinetic theory point of view, introducing the *one-particle* distribution function as a solution of the Vlasov equation [7,8] and the associated averaged Vlasov equation, that we will introduce later. Our final result in this direction is contained in *theorem 5.3.7*, which can be stated in words in the following way:

For narrow distribution functions and in the ultra-relativistic regime, one can use the charged cold fluid model as a good approximation to the Vlasov model, in the dynamical description of ultra-relativistic bunches of charged particles. The error of the approximation is of the same order as the area of the support of the distribution function in velocity space.

What this result means is that, in this regime, if the Vlasov equation holds, the differential equation defining the charged cold fluid equation holds approximately. Also, we note that a precise statement is involved, requiring some technical assumptions which we will discuss in the appropriate place.

In order to achieve the above results, a connection associated with the Lorentz force connection will be introduced (the Lorentz connection). We introduce an averaged version of this connection (the averaged Lorentz connection). The main advantage of this technique is that the averaged connection is simpler than the original one. This allows us to perform calculations whose results are not easily obtainable using to do in any other way.

1.1.1 Other results of the thesis

Another application of the theory of the averaged Lorentz model is the following. After introducing the Jacobi equation of an affine connection, we discuss the Jacobi equation associated with the averaged Lorentz connection. Then we prove that the linear dynamics used in accelerator physics [9-11] is an approximation to the Jacobi equation of the averaged Lorentz connection. Based on this interpretation, we define a notion of reference trajectory in beam dynamics in terms of observable quantities. Also, using the averaged Lorentz equation, we provide observable consequences of the collective nature of the bunch of particles.

We have also considered the question of how the gauge invariance principle affects the interpretation of the Lorentz force equation as an Euler-Lagrange equation of a functional. This led us to a precise definition of semi-Randers space.

Finally, we should mention that during the analysis of the main problem considered in this thesis and its mathematical formalization we found a generalization of

the notion of *connection* in differential geometry. We call this object *almost (projective) connection*. This is described in the *appendix*.

1.2 Structure of the thesis

In this *chapter* we introduce some notation and conventions that we will follow through this thesis.

In *chapter 2*, an introduction to relativistic kinetic theory is provided, following reference [7]. Then we define the charged cold fluid model and consider the asymptotic method developed in [6]. We will state the main problem considered in this thesis and give a short out-line of the strategy to solve it.

In *chapter 3*, we introduce the theory of non-linear connections defined by a second order differential system [12,13]. We also introduce the formalism and the notion of averaged connection, following the method already contained in ref. [22]. In particular we define the average of linear connections acting on sections of some relevant bundles (the pull-back bundles $\pi^*\mathbf{T}^{(p,q)}\mathbf{M}$). The data that we need to determine these connections is a system of second order differential equations called semi-spray. Those connections are obtained basically from the structure of the corresponding differential equations.

The original content of the Thesis constitutes chapters 4, 5, 6, 7 and appendix 4.

In *chapter 4*, the notion of semi-Randers space is introduced and discussed as a geometric description of the interaction of charged point particles with an external electromagnetic field [14, 15]. Then the Lorentz connection is obtained. Following the theory described in *chapter 3*, the corresponding averaged Lorentz connection is determined. It turns out that, if the dynamics happen in the ultra-relativistic limit and the support of the probability distribution function f is narrow (in a sense to be specified), the solutions of the Lorentz force equation can be approximated by the solutions of the averaged Lorentz force equation. We give an estimate of the

approximation as a function of the time of evolution, the energy of the system and the diameter of the distribution.

In *chapter 5* it is proved that under the same assumptions as in *theorem 4.6.6*, the relativistic charged cold fluid model can be obtained as an approximation from a kinetic model. The method that we follow to obtain this conclusion is the following. First, we introduce the averaged Vlasov equation and compare it with the original Vlasov equation. In particular we prove that, for the same initial conditions, both models have similar solutions in the ultra-relativistic regime when the distributions functions are narrow. After this, we use the averaged Vlasov model to give a bound on the acceleration of the main velocity vector field of the averaged Vlasov model. This bound is given in terms of the diameter of the distribution function, the energy and the time evolution. Then we prove that the mean velocity field associated with the solution of the Vlasov equation is similar to the mean velocity field obtained from the solution of the averaged Vlasov equation. This fact finishes the proof of *theorem 5.3.7*, which is the answer to the main problem considered in this thesis.

In *chapter 6* we use the Jacobi equation of the averaged connection to provide a geometric formulation of the transverse and longitudinal linear beam dynamics. Particular examples illustrate the general formalism. We obtain from the Jacobi equation of the averaged Lorentz connection the equations of motion of the transverse dynamics in magnetic dipole and quadrupole fields. In a similar way, the longitudinal dynamics in a constant and alternating electric field are obtained from the Jacobi equation. Although these are known examples, they illustrate the usefulness of the Jacobi equation of the averaged connection in beam dynamics. Corrections to the ordinary dynamics coming from *collective effects* are considered. As an application of the formalism we can provide a definition of *reference trajectory* that by construction is given in terms of observable quantities.

In *chapter 7* we discuss some of the results presented in this thesis as well as perspectives for further developments.

The present thesis work has produced the following articles and pre-prints [14], [15] and [16]:

1. R. Gallego Torromé, *On the Notion of Semi-Randers Spaces*, arXiv:0906.1940.
2. R. Gallego Torromé, *Geometric Formulation of the Classical Dynamics of Charged Particles in a External Electromagnetic Field*, arXiv:0905.2060, submitted.
3. R. Gallego Torromé, *Fluid Models from Kinetic Theory using Geometric Averaging*, arXiv:0912.2767, submitted.
4. R. Gallego Torromé, *Averaged Lorentz Dynamics and an Application in Plasma Dynamics*, arXiv:0912.0183, accepted to publish in the Proceedings of the XVIII Fall Meeting in Geometry and Physics, American Physics Society.

1.3 General conventions used in the thesis

For the main physical applications in this thesis, the space-time structure will be a flat four dimensional manifold endowed with a Lorentzian metric η with signature $(+, -, -, -)$. However, some results and techniques are valid for arbitrary dimension, signature and curvature. In these cases, it is explicitly stated. Sometimes we will require, to simplify the calculations, that the metric η is flat. When this is the case, it will be indicated. In all cases we assume that the space-time manifold \mathbf{M} is time-oriented.

Einstein summation convention is considered for any identical and repeated covariant and contravariant indices, if the contrary is not stated. All Latin indices run from 0 to $n - 1$, where n is the dimension of the space-time manifold. Vector notation is used for the spatial components (with respect to a given frame) of a vector. Indices are lowered using the metric η_{ij} and raised using the inverse metric η^{ij} , unless anything else is stated. The exterior product and the exterior derivative are normalized as in reference [17].

We have adopted the following convention for the physical parameters and con-

stants appearing in the models,

$$q = 1, m = 1, \epsilon_0 = 1, \mu_0 = 1, c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} = 1,$$

where ϵ_0 μ_0 are the dielectric and magnetic permeability constants of the vacuum; m is the mass and q the charge of the species of particles that we are considering.

We also use the following convention [3, pg 618]:

$$\vec{D} = \vec{E} + \vec{P}, \quad \vec{H} = \vec{B} - \vec{M}. \quad (1.3.1)$$

\vec{P} is the polarization vector and \vec{M} is the magnetization of the medium. Since we are considering that the bunch of particle propagates in the vacuum, we have that $\vec{P} = 0$, $\vec{M} = 0$. Hence, one has $\vec{E} = \vec{D}$, $\vec{H} = \vec{B}$.

The Maxwell equations are

$$\vec{\nabla} \cdot \vec{E} = \rho, \quad \vec{\nabla} \times \vec{B} = (\vec{J} + \frac{\partial \vec{E}}{\partial t}), \quad (1.3.2)$$

$$\vec{\nabla} \cdot \vec{B} = 0, \quad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \quad (1.3.3)$$

The Maxwell equations can also be written in a covariant form in the following way:

$$\partial_i \mathbf{F}_{jk} + \partial_k \mathbf{F}_{ji} + \partial_j \mathbf{F}_{ik} = 0, \quad {}^\eta \nabla_i \mathbf{F}^i{}_j = \eta_{kj} J^k, \quad (1.3.4)$$

where ${}^\eta \nabla_i$ is the covariant derivative associated to the Levi-Civita connection along the direction e_i .

The electromagnetic tensor is described by a 2-form, that in a local frame determines the following matrix:

$$\mathbf{F}_{ij}(x) = \begin{pmatrix} 0 & E_1(x) & E_2(x) & E_3(x) \\ -E_1(x) & 0 & -B_3(x) & B_2(x) \\ -E_2(x) & B_3(x) & 0 & -B_1(x) \\ -E_3(x) & -B_2(x) & B_1(x) & 0 \end{pmatrix}.$$

The Lorentz force is written as

$$\vec{F} := q(\vec{E} + \vec{v} \times \vec{B}), \quad \vec{v} = \frac{d\vec{\sigma}}{dt}, \quad \vec{F} = m \frac{d(\gamma \vec{v})}{dt}. \quad (1.3.5)$$

The parameter τ is the proper-time along σ associated with the metric η . In covariant formalism, the Electromagnetic field is described by a 2-form $\mathbf{F} = \mathbf{F}_{ij} dx^i \wedge dx^j$.

The Lorentz force equation for a mass $m = 1$ and a charge $q = -1$ is

$$\frac{d^2 \sigma^i}{d\tau^2} = -\mathbf{F}^i{}_j \frac{d\sigma^j}{d\tau}. \quad (1.3.6)$$

There are several categories of metric structures where we will work. The most general is semi-Riemannian category [47]. The results stated in this category refer to structures $\{\eta(x)\}$ which are non-degenerate, symmetric bilinear forms for each fixed $x \in \mathbf{M}$, and are smoothly defined on the n -dimensional manifold \mathbf{M} .

The second most general category refers to Lorentzian manifolds. In this case the metrics $\{\eta(x)\}$ are bilinear, symmetric forms with signature $(+, -, -, -)$ for fixed $x \in \mathbf{M}$ and are smoothly defined on the manifold is four dimensional space-time because physical reasons. Those results be defined may also be defined for n -dimensional space-times with signature $(+, -, \dots, -)$.

The third category of geometry is Minkowski geometry. In this case the metric is the Minkowski metric but the manifold considered is a domain inside the space \mathbf{R}^4 . In this category it makes sense to speak of global inertial frames: a global inertial frame (e_0, e_1, e_2, e_3) on \mathbf{M} such that the metric η has the following metric components:

$$(\eta_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The contraction operation is defined on the following way: given a tangent vector $W \in \mathbf{T}_x \mathbf{M}$, there is an homomorphism on the space of covariant tensors over x

denoted by:

$$\begin{aligned}\iota_W : \mathbf{T}_x^{(0,p)}\mathbf{M} &\longrightarrow \mathbf{T}_x^{(0,p-1)}\mathbf{M} \\ T &\mapsto \iota_W T\end{aligned}$$

given by

$$\iota_W T(X_1, \dots, X_{p-1}) := T(W, X_1, \dots, X_p), \quad \forall X_i \in \mathbf{T}_x\mathbf{M}, \quad W \in \mathbf{T}_x\mathbf{M}.$$

A similar definition applies pointwise to sections of $\mathcal{F}(M)$ -multilinear maps of vector fields contracted with a given vector field given by

$$\iota_W T(X_1, \dots, X_{p-1})(x) := T(x)(W, X_1, \dots, X_p), \quad \forall X_i \in \Gamma(\mathbf{TM}), \quad W \in \Gamma(\mathbf{TM}).$$

When we write down the results, we try to formalize in the largest category possible. Generally speaking, results from *chapter 3* fall into the category of semi-Riemannian metrics (indeed, some of them are even more generic than for metric structures). Results in *chapter 4* fall into this category of semi-Riemannian category, except for the main comparison results, where we explicitly use the flatness property of the metric η in some of the calculations.

In *chapter 5*, the results depend on the results of *chapter 4*. Therefore, although some of them are formulated for semi-Riemannian manifolds, the main results are formulated for compact domains of the Minkowski space.

In *chapter 6*, the main results are formulated for Minkowski space, since we have in mind to apply the geometric formalism to describe the behavior of beams of particles in accelerators, where gravitational effects are usually neglected.

In *chapter 7*, we point out the general conclusions of this thesis as well as open problems proposed.

Chapter 2

Fluid and kinetic models for ultra-relativistic beams

In this chapter we consider some basic notions that we will use later. In the same way, we introduce additional notation. There is also a short introduction to fluid models, kinetic models and to the asymptotic model described in [6].

2.1 Basic relativistic kinetic theory

In this *section* we review some elementary notions of the covariant kinetic theory which are relevant for our work. We mainly follow the notation of reference [7]. We will consider collision-less processes and detailed balance processes.

2.1.1 Intrinsic covariant formalism for relativistic kinetic theory

In this thesis, the kinetic models are based on the following general assumptions:

1. The space-time manifold \mathbf{M} is 4-dimensional and it is endowed with a Lorentzian metric η . The signature of the metric is $(1, -1, -1, -1)$, and the space-time is time orientable [7] and oriented. In general, the metric η is not flat. However we will use the flatness condition on the metric η in some of the calculations.

The Lorentzian metric η has an associated Levi-Civita connection ${}^\eta\nabla$. Also, η determines the Hodge star operator

$$\star : \Gamma \wedge^p \mathbf{M} \longrightarrow \Gamma \wedge^{4-p} \mathbf{M}$$

$$\omega_{i_1 \dots i_p} e^{i_1} \wedge \dots \wedge e^{i_p} \mapsto \omega_{i_1 \dots i_p} e^{i_{p+1}} \wedge \dots \wedge e^{i_{4-p}} \epsilon^{i_1 \dots i_p}_{i_{p+1} \dots i_{4-p}}.$$

$\Gamma \wedge^p \mathbf{M} := \{\omega : \mathbf{M} \rightarrow \wedge^p \mathbf{M}\}$ is the set of smooth sections of the vector bundle $\wedge^p \mathbf{M} \rightarrow \mathbf{M}$, with $\wedge^p \mathbf{M}$ the bundle of smooth p -forms over \mathbf{M} ; $\epsilon^{i_1 \dots i_p}_{i_{p+1} \dots i_{4-p}}$ is the total skew-symmetric symbol, where the indices are raised using the Lorentzian metric η and with $\epsilon_{0123} = 1$. $\omega_{i_1 \dots i_p} e^{i_1} \wedge \dots \wedge e^{i_p}$ is an arbitrary p -form expressed in a dual basis $\{e^0, \dots, e^3\}$ of an orthonormal basis $\{e_0, \dots, e_3\}$. The dual of a vector field is the 1-form defined pointwise by the relation

$$(V)^\flat(x) = \eta(V(x), \cdot).$$

Similarly, the dual of a 1-form is a vector field defined pointwise by the relation

$$\omega^\sharp(x) = \eta^{-1}(\omega, \cdot),$$

where η^{-1} is the bilinear form

$$\eta^{-1} : \mathbf{T}^* \mathbf{M} \times \mathbf{T}^* \mathbf{M} \longrightarrow \mathbf{R}$$

$$(\omega, \phi) \mapsto \eta^{-1}(\omega, \phi) := \eta(\omega^\sharp, \phi^\sharp).$$

2. The electromagnetic field is encoded in the 2-form \mathbf{F} , while the excitation field is encoded in the 2-form \mathbf{G} and the current density is a 3-form \mathbf{J} , all living on \mathbf{M} . They satisfy Maxwell's equations, which can be written in terms of differential forms as

$$d\mathbf{F} = 0, \quad d \star \mathbf{G} = \mathbf{J}; \quad (2.1.1)$$

$d : \wedge^p \mathbf{M} \longrightarrow \wedge^{p+1} \mathbf{M}$ is the exterior derivative operator acting on forms. This is a coordinate free form of the equations (1.3.4).

3. The relation between \mathbf{F} and \mathbf{G} is given by the constitutive relations. We assume that these relations are linear and in particular we put $\mathbf{F} = \mathbf{G}$, since the electromagnetic medium that we are considering is the vacuum.
4. The matter content of the models consists of a collection of identical charged point particles. The trajectory $\sigma(s)$ of each particle follows the Lorentz force equation

$$\frac{dy^i}{d\tau} = \mathbf{F}^i{}_j(\sigma(\tau)) y^j, \quad y^j = \frac{d\sigma^j(\tau)}{d\tau}, \quad (2.1.2)$$

where τ is the proper-time associated with the trajectory. The parameter τ is such that $\eta(\frac{d\sigma(\tau)}{d\tau}, \frac{d\sigma(\tau)}{d\tau}) = 1$.

5. The support of the one-particle distribution function $f(x, y)$ is in the 7-dimensional unit hyperboloid bundle,

$$\Sigma := \{(x, y), x \in \mathbf{M}, y \in \mathbf{T}_x\mathbf{M}, \eta(y, y) = 1, y^0 > 0\}. \quad (2.1.3)$$

A global coordinate system on Σ is (x^a, y^i) , $a = 0, 1, 2, 3$, $i = 1, 2, 3$, induced from any natural coordinate system on \mathbf{TM} . In the unit hyperboloid, y^0 is given as a function of y^1, y^2, y^3 and x^a . The manifold $\Sigma_x := \{y \in \mathbf{T}_x\mathbf{M}, \eta(y, y) = 1, y^0 > 0\}$ is called the unit hyperboloid over x .

6. There is defined a volume form on the unit hyperboloid bundle Σ . This volume form is obtained in terms of the metric η . On the tangent space \mathbf{TM} there is a volume 8-form

$$\sqrt{|det\eta|} dy^0 \wedge \dots \wedge dy^3 \wedge dx^0 \wedge \dots \wedge dx^3, \quad d^4(x) = dx^0 \wedge \dots \wedge dx^3,$$

with $\sqrt{\eta}$ the determinant of the matrix associated to the metric in a given coordinate system. The isometric embedding $e : \Sigma \hookrightarrow \mathbf{M}$ induces a volume form on the manifold Σ . We denote this volume form by $dvol(x, y) \wedge d^4x$. Since the space-time manifold \mathbf{M} is 4-dimensional, the volume form $dvol(x, y) \wedge d^4x$ is a 7-form.

The volume form $dvol(x, y)$ on Σ_x is obtained contraction of $dvol(x, y) \wedge d^4x$

on the orthogonal frame $\{e_0, \dots, e_3\}$: $d^4x(e_0, \dots, e_3) = 1$.

7. The Liouville vector field ${}^L\chi$ is tangent to the hyperboloid Σ . Using the conventions of *section 1.3*, the Liouville vector field ${}^L\chi$ can be written using local coordinates as

$${}^L\chi = y^i \partial_i + (F^i{}_j y^j - \Gamma^i{}_{jk} y^j y^k) \frac{\partial}{\partial y^i}, \quad i, j, k = 0, 1, 2, 3. \quad (2.1.4)$$

Remark Note that we have adopted the *extrinsic* formalism, where y^0 is considered an independent coordinate. Later we will explain the relation between the intrinsic and extrinsic formalism, and that they are equivalent for our purposes.

8. The one-particle distribution function $f(x, y)$ is defined over Σ and satisfies the equation

$${}^L\chi(f) = 0. \quad (2.1.5)$$

Equation (2.1.5) corresponds to the Vlasov equation in plasma physics and kinetic theory. The one-particle distribution function $f(x, y)$ is introduced as the probability density of finding a particle at the point $x \in \mathbf{M}$ with velocity vector $y \in \mathbf{T}_x \mathbf{M}$ [7]. This interpretation was supported in [7] using balance arguments and assuming that $f(x, y)$ is continuous. We will also assume additional smoothness and regularity conditions for $f(x, y)$.

2.1.2 Extrinsic formulation of the kinetic model

There is an alternative description of a kinetic model to the intrinsic one. In this alternative description the calculations are performed on the whole tensor bundle \mathbf{TM} and then the results are restricted to the unit hyperboloid. One uses the constraint $\eta(y, y) = 1$ when it is necessary. Note that the action of ${}^L\chi$ is on the ring of smooth functions of the hyperboloid Σ , since ${}^L\chi$ is a tangent vector to the unit hyperboloid Σ . This follows from the fact that ${}^L\chi|_{\Sigma} \cdot (\eta(y, y)) = 0$ and that the function $\eta(y, y) = \eta_{ij}(x) y^i y^j$ generates a foliation of \mathbf{TM} . A formal proof of this

fact can be found for instance in [18].

We will also define later an averaged Liouville vector field $\langle {}^L\chi \rangle$. This vector field lives on \mathbf{M} rather than on the tangent bundle \mathbf{TM} . However, it defines a second order differential equation and therefore a Liouville equation. The flow of $\langle {}^L\chi \rangle$ does not preserve the function $\eta(y, y)(x) := \eta_{ij}(x)y^i y^j$. Indeed, it is not guaranteed that the flow will preserve a *structure*. Therefore for the study of these kind of flows, it is more convenient to adopt the external formalism.

Using the volume form $dvol(x, y)$ one can obtain the velocity moments of the distribution $f(x, y)$. Therefore one can define moments of the distribution function, which are the expectation values of polynomials on y . With these moments, one can define the mean velocity field, the covariant kinetic energy-momentum tensor and the covariant energy-momentum flux tensor:

$$V^i(x) = \frac{1}{\int_{\Sigma_x} f(x, y) dvol(x, y)} \int_{\Sigma_x} y^i f(x, y) dvol(x, y). \quad (2.1.6)$$

$$T^{ij}(x) = \frac{1}{\int_{\Sigma_x} f(x, y) dvol(x, y)} \int_{\Sigma_x} y^i y^j f(x, y) dvol(x, y), \quad (2.1.7)$$

$$Q^{ijk}(x) = \frac{1}{\int_{\Sigma_x} f(x, y) dvol(x, y)} \int_{\Sigma_x} y^i y^j y^k f(x, y) dvol(x, y), \quad (2.1.8)$$

The balance equation for the number of particles implies the relations [7]

$$\eta \nabla_i V^i(x) = \frac{1}{\int_{\Sigma_x} f(x, y) dvol(x, y)} \int_{\Sigma_x} {}^L\chi(f) dvol(x, y), \quad (2.1.9)$$

$$\eta \nabla_j T^{ij}(x) = \mathbf{F}^i{}_j V^j + \frac{1}{\int_{\Sigma_x} f(x, y) dvol(x, y)} \int_{\Sigma_x} y^i {}^L\chi(f) dvol(x, y). \quad (2.1.10)$$

Since f follows the Liouville equation (2.1.5), one obtains

$$\eta \nabla_i V^i(x) = 0, \quad \eta \nabla_j T^{ij}(x) = \mathbf{F}^i{}_j V^j.$$

2.2 Relativistic charged cold fluid model

We introduce some geometric and physical objects that we need in the description of the asymptotic expansion of the relativistic cold fluid model proposed in [6]. The electromagnetic field is encoded in the 2-form \mathbf{F} , which is a solution of the Maxwell equations (2.1.1). The external electromagnetic field \mathbf{F} is created by the external current density \mathbf{J} such that in the space time regions that we will consider, one has that $\mathbf{J}(x) = 0$. The current density \mathcal{J} describes a system of charged point particles which also contributes to the total electromagnetic field. The whole dynamics is non-linear and one needs additional information to completely determine the dynamics. There are two additional pieces of information:

1. One has to postulate the dynamic equation for the current density \mathcal{J} . Examples for these joint dynamics are the Maxwell-Lorentz system, Maxwell-Vlasov, Klimontovich-Maxwell's system [8, *section 2.5*] and Maxwell-Lorentz-Dirac system [5].
2. In order to completely determine the system, constitutive relations between \mathbf{F} and \mathbf{G} are needed. We assume that the constitutive relations are $\mathbf{G} = \epsilon_0 \mathbf{F}$. We have adopted units such that $\epsilon_0 = 1$.

In flat regions, the metric η admits a set of translational Killing vectors $\{K_i, i = 0, 1, 2, 3\}$, $\mathcal{L}_{K_i} \eta = 0$.

Using differential forms, one can write conservation laws in a geometric way. For any vector field W on \mathbf{M} there is an associated *drive 3-form* [17]:

$$\tau_W^{em} = \frac{1}{2}(\iota_W \mathbf{F} \wedge \star \mathbf{F} - \iota_W \star \mathbf{F} \wedge \mathbf{F}).$$

For the case of Killing vector fields, the exterior derivative of τ_W^{em} is

$$d\tau_W^{em} = -\iota_W \mathbf{F} \wedge \mathbf{J},$$

where $\iota_W \mathbf{F}$ is the contraction of the vector field W with the 2-form \mathbf{F} . In the region outside of the sources $\mathbf{J} = 0$,

$$d\tau_W^{em} = 0. \quad (2.2.1)$$

In the presence of matter, equation (2.2.1) has to be generalized. For example, let us consider a model for matter described by a time-like vector field V . Then for a *dust*, the stress-energy tensor is

$$T(x) = \mathcal{N} V^\flat(x) \otimes V^\flat(x). \quad (2.2.2)$$

\mathcal{N} is a regular scalar density field and the velocity field is normalized, $\eta(V, V) = 1$. The current density \mathcal{J} is proportional to the velocity field:

$$\mathcal{J} = \mathcal{N} \star (V)^\flat. \quad (2.2.3)$$

and then $d\mathcal{J} = 0$. Combined with the assumption of the total momentum conservation $d(\tau_W^{em} + \star \iota_W T) = 0$ implies the field equation of motion for the fluid:

$$\eta \nabla_V V(x) = (\iota_V \mathbf{F})^\sharp \quad (2.2.4)$$

as a balance equation [17, pg 242-243], [19].

The dynamics of the relativistic charged cold fluid model is described by the following coupled system of differential equations,

$$d\mathbf{F} = 0, \quad d \star \mathbf{F} = -\rho \star V^\flat, \quad \eta \nabla_V V(x) = (\iota_V \mathbf{F})^\sharp, \quad \eta(V, V) = 1. \quad (2.2.5)$$

Although this is a complete model, in this thesis we will work with external electromagnetic fields. In this context, the mathematical and physical analysis are highly simplified.

2.3 Asymptotic expansion of the relativistic charged cold fluid model

Let us consider the following 1-parameter family of differential forms and vector fields:

$$V^\epsilon = \sum_{n=-1}^{+\infty} \epsilon^n V_n, \quad \rho^\epsilon = \sum_{n=1}^{+\infty} \epsilon^n \rho_n, \quad \mathbf{F}^\epsilon = \sum_{n=-1}^{+\infty} \epsilon^n \mathbf{F}_n, \quad (2.3.1)$$

where

$$V_n \in \Gamma \mathbf{TM}, \quad \rho_n \in \Gamma \wedge^0 \mathbf{M}, \quad \mathbf{F}_n \in \Gamma \wedge^2 \mathbf{M} \quad (2.3.2)$$

and ϵ is a small parameter. Substituting these expansions in equation (2.2.5) and equating terms of equal in ϵ , one obtains enough conditions to determine the fields (2.3.2) inductively [6]. For instance, the leading order terms are the vector field V_{-1} and the 2-form \mathbf{F}_{-1} such that:

$${}^\eta \nabla_{V_{-1}} V_{-1} = \iota_{V_{-1}} \mathbf{F}_{-1}, \quad d\mathbf{F}_{-1} = 0, \quad d \star \mathbf{F} = 0, \quad \eta(V_{-1}, V_{-1}) = 0.$$

Given initial data for V_{-1} and \mathbf{F}_{-1} , these equations are compatible. Note that they describe a charged mass-less fluid interacting with an external electromagnetic fluid.

In general the equations for the fields appearing in the expansion (2.3.1) are obtained from the equations (2.2.5). The procedure for the higher orders is as follows:

1. Consider a given electromagnetic field \mathbf{F}_{-1} , solution of the differential equations

$$d\mathbf{F}_{-1} = 0, \quad d \star \mathbf{F}_{-1} = 0 \quad (2.3.3)$$

for some initial value of \mathbf{F}_{-1} on a space-like hypersurface. Physically \mathbf{F}_{-1} is interpreted as the external electromagnetic field.

2. Then one has to solve the equation

$${}^\eta \nabla_{V_{-1}} V_{-1} = (\iota_{V_{-1}} \mathbf{F}_{-1})^\sharp \quad (2.3.4)$$

subject to the condition

$$\eta(V_{-1}, V_{-1}) = 0$$

and for given initial data in the space-like hypersurface $t = t_0$. This is possible because equation (2.3.4) can be re-written as an ordinary differential equation and one can apply standard results (see for instance the *appendix*).

3. Then one solves the equation for ρ_1 , which is

$$d \star (\rho_1 V_{-1}^b) = 0, \quad (2.3.5)$$

for given initial values in a space-like hypersurface.

4. The 2-form \mathbf{F}_0 is a solution to the Maxwell equation, that can be written as

$$d\mathbf{F}_0 = 0, \quad d \star \mathbf{F}_0 = - \star \rho_1 V_{-1}^b, \quad (2.3.6)$$

where one has to specify the initial values on a space-like hypersurface.

5. V_0 is the solution of the equation

$$\eta \nabla_{V_{-1}} V_0 + \eta \nabla_{V_0} V_{-1} = (\iota_{V_{-1}} \mathbf{F}_0 + \iota_{V_0} \mathbf{F}_{-1})^\sharp \quad (2.3.7)$$

subject to the requirement that $\eta(V_{-1}, V_0) = 0$ and for fixed initial values of the vector field V_0 in a space-like hypersurface.

6. The density ρ_2 is defined as the solution of

$$d \star (\rho_2 V_{-1}^b) + d \star (\rho_2 V_{-1}^b) = 0, \quad (2.3.8)$$

again after given the initial values of ρ_2 on a space-like hypersurface.

7. The equation

$$d \star \mathbf{F}_1 = - \star \rho_2 V_{-1}^b - \star \rho_1 V_0^b \quad (2.3.9)$$

can be solved for \mathbf{F}_1 , once the initial values for \mathbf{F}_1 are specified.

8. V_1 is a solution of the equation

$${}^\eta\nabla_{V_{-1}}V_1 + {}^\eta\nabla_{V_0}V_{-1} + {}^\eta\nabla_{V_1}V_{-1} = (\iota_{V_{-1}}\mathbf{F}_0 + \iota_{V_0}\mathbf{F}_{-1} + \iota_{V_1}\mathbf{V}_{-1})^\sharp. \quad (2.3.10)$$

We need to specify the initial values on a space-like hypersurface.

Through a generalization of this procedure, the fields (2.3.2) can be solved order by order in ϵ . The only non-linear differential equation to be solved is for V_{-1} . Indeed it can be written as an ordinary second order differential equation for the integral curves of V_{-1} . These properties make it easier to solve both the analytical and numerical treatment of the problem than to solve the original equations (2.2.5), which are a system of non-linear and coupled partial differential equations.

The vector field V_{-1} has a difficult physical interpretation, because it corresponds to a charged cold fluid composed of mass-less particles and at the same time interacting with an external electromagnetic field. There is no known classical physical system (in vacuum) with such characteristics (in quantum physics, the low energy limit of graphene admits states which are mass-less and interact with the electromagnetic field [56]).

2.4 Statement of the main problem considered in this thesis and out-line of the strategy to solve it

It was claimed that the model introduced in [6] is able to provide a consistent treatment of the back reaction and self-force problems that appear in classical electrodynamics, for some situations which are of practical interest like ultra-relativistic plasmas. This claim is based on the assumption that the charged cold fluid model is an acceptable description of the dynamics of bunches of particles in the ultra-relativistic regime.

On the other hand, fluid models have been used intensively in the description of the dynamics of plasmas [8, 37-39]. However, these usual models are based on assumptions on the moments of the distribution function, which are difficult to check

in experimental conditions.

These motivate the main problem considered in this thesis:

Is it mathematical justified to use the charged cold fluid model (2.2.5) under the conditions present in the currently used particle accelerators?

We will estimate the value of the differential operators appearing in the equation $\eta \nabla_V V(x) = (\iota_V \mathbf{F})^\sharp$, where V is the mean field (2.1.6) for a given distribution function f . Then we will show in *chapter 5* that the differential expression for the charged cold fluid model equation is bounded and controlled by powers of the diameter α of the distribution function $f(x, y)$, powers of the *energy*¹ of the system and powers of the coordinate time evolution t . The relation is such that for narrow distribution functions and in the ultra-relativistic regime, the charged cold fluid model is a good approximation of the kinetic model.

The strategy that we will follow is the following. Since the fluid model $V(x)$ is an approximate description of the system, we interpret $V(x)$ as an averaged quantity. On the other hand, given a dynamical system, we can associate a *non-linear connection*. That connection can be averaged, using a distribution function. If in addition, the difference between the original connection and the averaged connection is small, one can substitute the original one by the averaged connection in the description of the dynamics.

Any dynamical system described by a connection has an associated kinetic model. In particular, this is true for the Lorentz force equation and the averaged Lorentz force equation, that we will define. Since the Lorentz connection is similar to the averaged Lorentz connection in a sense that we will explain later, one can also substitute the associated kinetic models.

Working with the averaged model has technical advantages. In particular one can give estimates of the value of some differential operators which appear in the fluid models. This will be done for the charged cold fluid model.

¹The notion of energy of a bunch of particles that we will use in those bounds is not trivial and will be introduced in *chapter 4*.

Chapter 3

The averaged connection

In this *chapter* we introduce the notions of non-linear connections and the associated averaged connections, before applying the method to the connection associated with the Lorentz force equation in the next *chapter*. The construction is adapted from reference [22]. This *chapter* explains the mathematical theory that we will use in *chapters 4, 5 and 6*.

3.1 Non-Linear connection associated with a second order differential equation

3.1.1 Second order differential equations and the associated non-linear Berwald-type connection

Let \mathbf{M} be an n -dimensional smooth manifold. A *natural coordinate system* on the tangent bundle $\pi : \mathbf{TM} \longrightarrow \mathbf{M}$ is constructed in the following way. Let (x, \mathbf{U}) be a local coordinate system on \mathbf{M} , where $\mathbf{U} \subset \mathbf{M}$ is an open sub-set of \mathbf{M} and $x : \mathbf{U} \rightarrow \mathbf{R}^n$ a local coordinate system. An arbitrary tangent vector at the point $p \in \mathbf{U}$ is of the form $X_p := X = X^k \frac{\partial}{\partial x^k} |_p$. The local coordinates associated with the tangent vector $X_p \in \mathbf{T}_x \mathbf{M} \subset \mathbf{TM}$ are (x^k, y^k) . We will identify the point $x \in \mathbf{M}$ with its coordinates, by notational convenience. \mathbf{N} is a sub-bundle of the tangent bundle \mathbf{TM} . From the imbedding $e : \mathbf{N} \hookrightarrow \mathbf{TM}$, $e(\mathbf{N})$ acquires the induced

differential structure from \mathbf{TM} ; $e(\mathbf{N})$ is denoted by \mathbf{N} .

We recall the following notion of connection [20, pg 314]. Let $\pi : \mathbf{N} \longrightarrow \mathbf{M}$ be a bundle over \mathbf{M} and consider the differential function $d\pi : \mathbf{TN} \longrightarrow \mathbf{TM}$. Then the vertical bundle is $\mathcal{V} = \ker(d\pi) \subset \mathbf{TN}$.

Definition 3.1.1 *A connection in the sense of Ehresmann is a distribution $\mathcal{H} \subset \mathbf{TN}$ such that*

1. *There is a decomposition at each point $u \in \mathbf{N}$, $\mathbf{T}_u\mathbf{N} = \mathcal{H}_u \oplus \mathcal{V}_u$.*
2. *The horizontal lift exists for any curve $t \mapsto \sigma(t) \in \mathbf{M}$, $t_1 \leq t \leq t_2$ and is defined for each $\xi \in \mathbf{T}_u\mathbf{M}$ and $u \in \pi^{-1}(x)$.*

Let us consider a set of n second order differential equations, with n the dimension of \mathbf{M} . The solutions are parameterized curves on \mathbf{M} . Assume that the system of differential equations describes the flow of a vector field ${}^G\chi \in \Gamma\mathbf{TN}$. In particular, the system of differential equations has the following form

$$\frac{d^2x^i}{dt^2} - G^i(x, \frac{dx}{dt}) = 0, \quad i = 1, \dots, n. \quad (3.1.1)$$

This system of differential equations is equivalent to the following system of first order differential equations on \mathbf{N} ,

$$\begin{cases} \frac{dy^i}{dt} - G^i(x, y) = 0, \\ \frac{dx^i}{dt} = y^i, \end{cases} \quad i = 1, \dots, n. \quad (3.1.2)$$

The coefficients $G^i(x, y)$ are called *spray coefficients* if they are homogeneous functions of degree one on the coordinate y ; in the general case where they are not homogeneous those coefficients are called semi-spray coefficients. $G^i(x, y)$ transform under a change of natural local coordinates on \mathbf{N} , induced from changes of coordinates on \mathbf{M} , in such a way that the system of differential equations (3.1.1) is *covariant*. Explicitly, if the change of local natural coordinates on the manifold \mathbf{N}

is

$$\begin{cases} \tilde{x}^i = \tilde{x}^i(x), \\ \tilde{y}^i = \frac{\partial \tilde{x}^i}{\partial x^j} y^j, \end{cases}$$

then the associated co-frame transforms as

$$\begin{cases} d\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^j} dx^j, \\ d\tilde{y}^i = \frac{\partial^2 \tilde{x}^i}{\partial x^k \partial x^j} y^k dx^j + \frac{\partial \tilde{x}^i}{\partial x^j} dy^j. \end{cases}$$

The induced transformation in the associated system of differential equations is

$$\frac{dy^i}{dt} - G^i(x, y) = 0 \Rightarrow \frac{d\tilde{y}^i}{dt} - \tilde{G}^i(\tilde{x}, \tilde{y}) = 0,$$

where the coefficients $\tilde{G}^i(x, y)$ are

$$\tilde{G}^i(\tilde{x}, \tilde{y}) = \sum_{j,k} \left(\frac{\partial \tilde{x}^l}{\partial x^j} \right) y^j \left(\frac{\partial \tilde{x}^s}{\partial x^k} \right) y^k \frac{\partial^2 \tilde{x}^i}{\partial x^l \partial x^s} - \sum_j \left(\frac{\partial \tilde{x}^i}{\partial x^j} \right) G^j(x, y).$$

The vertical distribution \mathcal{V} admits a local *holonomic* basis given by

$$\left\{ \frac{\partial}{\partial y^1}, \dots, \frac{\partial}{\partial y^n} \right\}, \quad i, j = 1, \dots, n. \quad (3.1.3)$$

Using these spray coefficients it is possible to define a horizontal n -dimensional distribution of the fiber bundle $\mathbf{TN} \longrightarrow \mathbf{N}$. The basis for the distribution is

$$\left\{ \frac{\delta}{\delta x^1}, \dots, \frac{\delta}{\delta x^n} \right\}, \quad \frac{\delta}{\delta x^k} := \frac{\partial}{\partial x^k} - \frac{\partial G^i}{\partial y^k} \frac{\partial}{\partial y^i}, \quad i, j = 1, \dots, n. \quad (3.1.4)$$

It generates a supplementary distribution to the vertical distribution. The non-linear connection coefficients $N^i{}_k(x, y)$ are defined by the relation

$$G^i(x, y) := y^k N^i{}_k(x, y).$$

For a spray, the connection coefficients of the non-linear connection are:

$$N^i{}_k(x, y) = \frac{\partial G^i(x, y)}{\partial y^k}.$$

Since the spray coefficients G^i are transformed under the a change in natural coordinates in a well defined way, the non-linear connection coefficients $N^i_j(x, y)$ are also transformed in a characteristic form [24, 34],

$$(\tilde{x}^i = \tilde{x}^i(x), \tilde{y}^i = \frac{\partial \tilde{x}^i}{\partial x^j} y^j) \Rightarrow \tilde{N}^i_m(x, y) \frac{\partial \tilde{x}^m}{\partial x^j}(x) = N^m_j(x, y) \frac{\partial \tilde{x}^i}{\partial x^m}(x) + \frac{\partial^2 \tilde{x}^i}{\partial x^k \partial x^j}(x) y^k.$$

Given a spray $G^i(x, y)$, we define the connection coefficients such that the only non-zero coefficients correspond to the covariant derivative of horizontal sections of $\Gamma\mathbf{TN}$ along horizontal sections of \mathbf{TN} and such that they are given by the Hessian of the spray:

$$\Gamma^i_{jk}(x, y) := \frac{1}{2} \frac{\partial^2 G^i(x, y)}{\partial y^j \partial y^k}.$$

All the other coefficients are zero. This type of connection resembles the so-called Berwald connection used in Finsler geometry [45]. From the point of view of the geometry of sprays, it is a natural connection. Note that, while $\Gamma^i_{jk}(x, y)$ can be associated with a linear connection on \mathbf{TN} , the connection coefficients $N^i_j(x, y)$ cannot (this is why they are called non-linear connection coefficients); $N^i_j(x, y)$ determines a connection which is non-linear in the direction of the derivation.

Given the non-linear connection, one can define the horizontal lift of the tangent vectors; the horizontal lift of $X = X^i \partial_i \in \mathbf{T}_x \mathbf{M}$ to the space $\mathbf{T}_u \mathbf{N}$ is defined by $h(X) = X^i \frac{\delta}{\delta x^i}$. This lift is defined here using local coordinates. However, an intrinsic definition can be found in [13]. After introducing this lift, one can define the horizontal lift of vector fields and tensor fields of the corresponding bundles.

3.1.2 The pull-back bundle

Let us consider the product $\mathbf{N} \times \mathbf{TM}$ and the canonical projections

$$\pi_1 : \pi^* \mathbf{TM} \longrightarrow \mathbf{N}, \quad (u, \xi) \longrightarrow u,$$

$$\pi_2 : \pi^* \mathbf{TM} \longrightarrow \mathbf{TM}, \quad (u, \xi) \longrightarrow \xi.$$

The pull-back bundle $\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$ of the bundle \mathbf{TM} is the *minimal* sub-bundle of the cartesian product $\mathbf{N} \times \mathbf{TM}$ such that the following equivalence relation holds: for every $u \in \mathbf{N}$ and $(u, \xi) \in \pi_1^{-1}(u)$, $(u, \xi) \in \pi^*\mathbf{TM}$ iff $\pi \circ \pi_2(u, \xi) = \pi(u)$; The pull-back bundle $\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$ is such that the following diagram commutes,

$$\begin{array}{ccc} \pi^*\mathbf{TM} & \xrightarrow{\pi_2} & \mathbf{TM} \\ \pi_1 \downarrow & & \downarrow \pi \\ \mathbf{N} & \xrightarrow{\pi} & \mathbf{M}. \end{array}$$

$\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$ is a real vector bundle with fibers diffeomorphic to $\mathbf{T}_x\mathbf{M}$. For instance, let $\{e_i, i = 0, \dots, n-1\}$ be a local frame for the sections of the tangent bundle \mathbf{TM} . Then $\{\pi^*e_i, i = 0, \dots, n-1\}$ is a local frame for the sections of the pullback bundle $\pi^*\mathbf{TM}$. Let $\xi^i(x, y)\pi_{(x, y)}^*e_i(x)$ be an arbitrary element in the fiber over $(x, y) \in \mathbf{N}$; the element $\pi_{(x, y)}^*e_i(x)$ is the unique element in $\pi^*\mathbf{TM}$ such that $(\pi \circ \pi_1)(\pi_{(x, y)}^*e_i(x)) = (\pi_2 \circ \pi)(\pi_{(x, y)}^*e_i(x))$ and that $\pi_1 \circ (\pi_{(x, y)}^*e_i(x)) = e_i(x)$.

Another way to visualize this pull-back bundle is the following. Let us consider the bundle $\pi : \mathbf{N} \longrightarrow \mathbf{M}$ and a fiber $\pi^{-1}(x) \subset \mathbf{N}$. On each point $u \in \pi^{-1}(x)$ we attach a copy of the vector space $\mathbf{T}_x\mathbf{M}$. This assignment is done by the definition of π^* on a local frame: $\pi^* : \{e_1(x), \dots, e_n(x)\} \longrightarrow \{\pi^*|_u e_1(x), \dots, \pi^*|_u e_n(x)\}$ and taking linear combinations of the elements of this local frame. When we consider sections of the bundle $\Gamma\pi^*\mathbf{TM}$ these linear combinations are u -dependent, instead of x -dependent.

Similarly, other pull-back bundles can be constructed from other tensor bundles over \mathbf{M} , for instance $\pi^*\mathbf{T}^*\mathbf{M} \longrightarrow \mathbf{N}$ and $\pi^*\mathbf{T}^{(p, q)}\mathbf{M} \longrightarrow \mathbf{N}$, with $\mathbf{N} \subset \mathbf{TM}$ a sub-bundle, $\mathbf{T}^*\mathbf{M}$ the vector bundle of 1-form over \mathbf{M} and $\mathbf{T}^{(p, q)}\mathbf{M}$ the bundle of (p, q) -tensors over \mathbf{M} .

Given a non-linear connection on the bundle $\mathbf{TN} \longrightarrow \mathbf{N}$, there are several related linear connections on the pull-back bundle $\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$.

Let χ a semi-spray defined on \mathbf{N} . We stipulate the following connection on $\pi^*\mathbf{TM}$,

defined by the conditions

$$\nabla_{\frac{\delta}{\delta x^j}} \pi^* Z := {}^x\Gamma(x, y)^i_{jk} Z^k \pi^* e_i, \quad \nabla_V \pi^* Z := 0, \quad V \in \mathcal{V}. \quad (3.1.5)$$

Here $\{\pi^* e_i, i = 0, \dots, n-1\}$ is a local frame for sections $\Gamma(\pi^* \mathbf{TM})$. This connection can be generalized to general tensor bundles over \mathbf{M} .

3.2 The average operator associated with a family of automorphisms

3.2.1 Average of a family of automorphisms

The averaged connection was introduced in the context of positive definite Finsler geometry in [22]. However, in this thesis we need to formulate the theory for arbitrary linear connections on the bundle $\pi^* \mathbf{TM} \rightarrow \mathbf{N}$, where $\mathbf{N} \rightarrow \mathbf{M}$ is a sub-bundle of the tangent bundle $\mathbf{TM} \rightarrow \mathbf{M}$.

Let π^*, π_1, π_2 be the canonical projections of the pull-back bundle $\pi^* \mathbf{T}^{(p,q)} \mathbf{M} \rightarrow \mathbf{N}$, $\mathbf{T}^{(p,q)} \mathbf{M}$ being the tensor bundle of type (p, q) over \mathbf{M} , $\pi_u^* \mathbf{T}^{(p,q)} \mathbf{M}$ the fiber over $u \in \mathbf{N}$ of $\pi^* \mathbf{T}^{(p,q)} \mathbf{M}$, $\mathbf{T}_x^{(p,q)} \mathbf{M}$ the tensor space over $x \in \mathbf{M}$, S_x a generic element of $\mathbf{T}_x^{(p,q)} \mathbf{M}$ and S_u is the evaluation of the section $S \in \Gamma(\pi^* \mathbf{T}^{(p,q)} \mathbf{M})$ at the point $u \in \mathbf{N}$.

For each tensor $S_z \in \mathbf{T}_z^{(p,q)} \mathbf{M}$ and $v \in \pi^{-1}(z)$, $z \in \mathbf{U} \subset \mathbf{M}$ the following isomorphisms are defined:

$$\pi_2|_v : \pi_v^* \mathbf{T}^{(p,q)} \mathbf{M} \longrightarrow \mathbf{T}_z^{(p,q)} \mathbf{M}, \quad S_v \mapsto S_z$$

$$\pi_v^* : \mathbf{T}_z^{(p,q)} \mathbf{M} \longrightarrow \pi_v^* \mathbf{T}_z^{(p,q)} \mathbf{M}, \quad S_z \mapsto \pi_v^* S_z.$$

To define the averaging operation we need two type of structures:

1. A family of non-intersecting, oriented sub-manifolds

$$\mathbf{N}_U := \bigsqcup_{x \in U} \mathbf{N}_x, \quad \mathbf{N}_x \subset \mathbf{T}_x \mathbf{M}.$$

2. A measure at each point $x \in \mathbf{M}$, which is an element $f(x, y) \omega_x(y) \in \bigwedge^m \mathbf{N}_x$, where m is the dimension of \mathbf{N}_x and $f_x : \mathbf{N}_x \rightarrow [0, \infty]$, $f_x := f(x, \cdot)$ is required to have compact support on \mathbf{N}_x .

Consider a family of endomorphisms, $\{A_u : \pi_u^* \mathbf{T}\mathbf{M} \rightarrow \pi_u^* \mathbf{T}\mathbf{M}, u \in \pi^{-1}(x)\}$. Let us consider the integral operations

$$\left(\int_{\mathbf{N}_x} \pi_2|_u A_u \pi_u^* \right) \cdot S(x) := \int_{\mathbf{N}_x} (\pi_2|_u A_u \pi_u^* S(x, u)) f(x, u) \omega_x(u),$$

The volume function is defined as

$$x \mapsto \text{vol}(\mathbf{N}_x) := \int_{\mathbf{N}_x} \omega_x(u) f(x, u).$$

Definition 3.2.1 Consider a family of endomorphisms,

$$\{A_w : \pi_w^* \mathbf{T}\mathbf{M} \rightarrow \pi_w^* \mathbf{T}\mathbf{M}, w \in \pi^{-1}(x)\}.$$

The average endomorphism of this family is the endomorphism:

$$\langle A \rangle_x : \mathbf{T}_x \mathbf{M} \rightarrow \mathbf{T}_x \mathbf{M}$$

$$S_x \mapsto \frac{1}{\text{vol}(\Sigma_x)} \left(\int_{\Sigma_x} \pi_2|_u A_u \pi_u^* \right) \cdot S_x,$$

$$u \in \pi^{-1}(x), S_x \in \Gamma_x \mathbf{M}.$$

We denote the averaged endomorphisms by symbols between brackets.

Remark. There is a similar notion which applies to families of homomorphisms, instead of endomorphisms between different vector bundles.

The averaging operation has the following effect. Let us consider an arbitrary

tensor S of a tangent space $\mathbf{T}_x^{(p,q)}\mathbf{M}$. Then the action of the integrand on S is obtained as follows:

1. First, $\pi_u^*S(x)$ moves S from the fiber $\pi^{-1}(x)$ to the fiber $\{\pi_1^{-1}(u), u \in \pi^{-1}(x) \subset \mathbf{N}\}$ of the bundle $\pi^*\mathbf{T}^{(p,q)}\mathbf{M}$.
2. On this image, the operator A_u acts: $A_u : \pi_1^{-1}(u) \longrightarrow \pi_1^{-1}(u)$.
3. The second projection again changes the fiber from $\pi_1^{-1}(u)$ to the fiber $\pi^{-1}(x)$. However, repeating this procedure for each $u \in \mathbf{U}$, being \mathbf{U} an open set. The result is not an element of $\mathbf{T}_x^{(p,q)}\mathbf{M}$, since there is a dependence on $u \in \Sigma_x$.
4. The integration of this variable provides the desired element, eliminating the dependence on u .

From this short discussion we observe that the geometric interpretation of the average operation is quite subtle. We are actually seeking an intrinsic definition, besides the general one in [50].

One can prove the following fact. The averaging operator acting on a element $S^i e_i$ is the following:

$$S^i(x)e_i(x) \mapsto \left(\int_{\Sigma_x} \omega_x(u) [A_u]^i_j S^j(x) \right) e_i(x),$$

where $[A_u]^i_j$ is the coordinate representation on a given basis of the linear operator A at the point $u \in \pi^{-1}(x)$.

3.2.2 Examples of geometric structures which provide an averaging procedure

1. *Lorentzian structures* [23]. The geometric data is a Lorentzian metric η defined on \mathbf{M} . The disjoint union of the family of sub-manifolds $\Sigma_x \subset \mathbf{T}_x\mathbf{M}$ defines the fibre bundle $\pi : \Sigma \longrightarrow \mathbf{M}$, which we called the unit hyperboloid bundle

over $x \in \mathbf{M}$,

$$\Sigma := \bigsqcup_{x \in \mathbf{M}} \{\Sigma_x \subset \mathbf{T}_x \mathbf{M}\}, \quad \Sigma_x := \{y \in \mathbf{T}_x \mathbf{M} \mid \eta(y, y) = 1\}$$

The manifold Σ_x is non-compact and oriented. The measure on Σ_x is given by the following $(n-1)$ -form

$$f(x, y) \omega_x(y) := f(x, y) \sqrt{\eta} \frac{1}{y^0} dy^1 \wedge \cdots \wedge dy^{n-1}, \quad y^0 = y^0(x^0, x^1, \dots, x^{n-1}, y^1, \dots, y^{n-1}),$$

The function y^0 defines the parameterized hypersurface $\Sigma_x \subset \mathbf{T}_x \mathbf{M}$, since y^0 can be solved from the condition $\eta_{ij(x)} y^i y^j = 1$. This equation can be expanded

$$\eta_{00} y^0 y^0 + 2 \sum_{a=1}^{n-1} \eta_{0a} y^0 y^a + \left(\sum_{a,b=1}^{n-1} \eta_{ab} y^a y^b \right) = 1.$$

There are two type of solutions for y^0 :

(a) If $\eta_{00} \neq 0$, one obtains a two-fold hyperboloid

$$y^0 = \frac{1}{\eta_{00}} \left(- \sum_{a=1}^{n-1} \eta_{0a} y^a \pm \sqrt{\left(\sum_{a=1}^{n-1} \eta_{0a} y^a \right)^2 - \eta_{00} \left(\sum_{a,b=1}^{n-1} \eta_{ab} y^a y^b - 1 \right)} \right).$$

(b) If $\eta_{00} = 0$, the solution for y^0 is:

$$y^0 = \frac{1 - \left(\sum_{a,b=1}^{n-1} \eta_{ab} y^a y^b \right)}{2 \sum_{a=1}^{n-1} \eta_{0a} y^a}.$$

Note that the Lorentzian metric η does not determine the manifolds $\{\Sigma_x, x \in \mathbf{M}\}$. For instance, one can consider $\tilde{\Sigma}$ to be the collection of null cones over \mathbf{M} :

$$\Sigma := \bigsqcup_{x \in \mathbf{M}} \{\mathbf{NC}_x \subset \mathbf{T}_x \mathbf{M}\}, \quad \mathbf{NC}_x := \{y \in \mathbf{T}_x \mathbf{M} \setminus \{0\} \mid \eta(y, y) = 0\}.$$

$\pi : \mathbf{NC} \longrightarrow \mathbf{M}$ is the null cone bundle over \mathbf{M} and \mathbf{NC}_x is the null cone over x ; on the other hand, $e : \mathbf{NC} \hookrightarrow \mathbf{TM}$ is a sub-bundle of $\mathbf{TM} \longrightarrow \mathbf{M}$.

2. *Finsler structures* [24, 45]. In this case, the Finsler function $F(x, y)$ defines the fundamental tensor $g_{ij}(x, y) = \frac{1}{2} \frac{\partial^2 F^2(x, y)}{\partial y^j \partial y^i}$ which is positive definite, homogeneous of degree zero on y , smooth and lives on the sub-bundle $\mathbf{N} := \mathbf{TM} \setminus \{0\}$. The bundle Σ is defined as the disjoint union,

$$\Sigma := \bigsqcup_{x \in \mathbf{M}} \{\mathbf{I}_x \subset \mathbf{T}_x \mathbf{M}\}, \quad \mathbf{I}_x := \{y \in \mathbf{T}_x \mathbf{M} \mid F(x, y) = 1\}.$$

Σ is the indicatrix bundle over \mathbf{M} . The manifold \mathbf{I}_x is compact and strictly convex for each $x \in \mathbf{M}$ and is the indicatrix at x . The volume form is

$$\omega_x(u) = d\text{vol}(x, y) := \sqrt{g} \frac{1}{y^0} dy^1 \wedge \cdots \wedge dy^{n-1}, \quad y^0 = y^0(y^1, \dots, y^{n-1}),$$

where the function y^0 is a solution of the implicit equation $F(x, y) = 1$. We can see that locally this equation has a solution using the implicit function theorem and the homogeneous properties of the function F (in particular, using Euler's theorem of homogeneous functions). If we take the derivative respect to y^0 of the function $\phi(x, y) = F(x, y) - 1$ and we put it equal to zero, we then get the condition of vanishing jacobian:

$$\frac{\partial}{\partial y^0}(F(x, y) - 1) = \frac{\partial}{\partial y^0}(g_{ij}(x, y)y^i y^j - 1) = 0.$$

Using Euler's theorem one obtains:

$$\begin{aligned} 0 &= \frac{\partial}{\partial y^0}(F(x, y) - 1) = (2g_{0j}(x, y)y^j + 2\frac{\partial}{\partial y^0}(g_{ij}(x, y))y^i y^j) = \\ &= (2g_{0j}(x, y)y^j + \frac{\partial}{\partial y^0}(\frac{\partial^2 F^2(x, y)}{\partial y^j \partial y^k})y^j y^k). \end{aligned}$$

Commuting the derivatives and considering Euler's theorem for homogenous functions, since F is homogeneous on y , the above expression is

$$= (2g_{0j}(x, y)y^j + y^j \frac{\partial}{\partial y^j}(\frac{\partial^2 F^2(x, y)}{\partial y^0 \partial y^k})y^k) = 2g_{0j}(x, y)y^j = 0.$$

Since the metric g is positive definite, the only solution is $y = 0$, which is

outside the indicatrix \mathbf{I}_x . Therefore, we can apply the hypothesis of the implicit function theorem and the equation $\phi(x, y) = 0$ can be solved for y^0 .

3. *Symplectic structures* [25]. In this case, there is defined on $\mathbf{T}^*\mathbf{M}$ a non-degenerate, closed 2-form ω . Due to Darboux's theorem [25, pg 246], there is a canonical local coordinate system of $\mathbf{T}^*\mathbf{M}$ such that the symplectic form ω can be written as

$$\omega = \sum_{i=0}^{n-1} dp_i \wedge dq^i.$$

Associated with ω there is defined on the dual tangent bundle $\mathbf{T}^*\mathbf{M}$ a volume $2n$ -form

$$S = \omega \wedge \cdots \wedge \omega.$$

Using canonical coordinates (q, p) , the $2n$ -differential form can be written as:

$$S(p, q) = dp^0 \wedge \cdots \wedge dp^{n-1} \wedge dq^0 \wedge \cdots \wedge dq^{n-1}.$$

Let us assume the existence of a nowhere zero vector field V on \mathbf{TM} (therefore the Euler characteristic of \mathbf{TM} must be different from zero). Then we can construct the $(2n - 1)$ -form

$$\omega_q(p) = \iota_V S, \quad V \in \Gamma\mathbf{T}(\mathbf{T}^*\mathbf{M})$$

$\iota_V S$ is a non-degenerate $(2n - 1)$ differential form whose value on V is zero, since $\iota_Z \iota_Z S = 0$ for any vector Z . Let us chose a distribution of commuting vector fields, $\{X_i, [X_i, X_j] = 0, \quad i, j = 1, \dots, 2n - 1\}$ locally supplementary to V such that $\{V, X_1, \dots, X_{2n-1}\}$ is a local frame of \mathbf{TM} . The distribution $\{X_1, \dots, X_{2n-1}\}$ is integrable and $\iota_V S$ is a volume form on the integral manifold \mathbf{S}^\perp . On the other hand \mathbf{S}^\perp is a fibered manifold:

$$\pi : \mathbf{S}^\perp := \bigsqcup_{x \in \mathbf{M}} \mathbf{S}^\perp_x \longrightarrow \mathbf{M}.$$

Therefore we can define an averaging operation.

An interesting thing about this example is that we can only construct *local averaging procedures*. The overlapping of open sets where the averaging operation is applied non-trivial and in general one needs more structures to define consistently the averaging procedure globally.

4. *Hermitian Vector Bundles* [25]. The construction is similar to the one in the *Finslerian case*. The sub-manifolds Σ_x are defined as:

$$\Sigma_x := \{y \in \mathbf{T}_x \mathbf{M} \mid H(y, y) = 1\},$$

where H is the hermitian structure on \mathbf{M} . Therefore

$$\Sigma := \bigsqcup_{x \in \mathbf{M}} \Sigma_x.$$

Let us assume that the hermitian structure is of the form $H = \eta + \imath\omega$, where η is a Riemannian structure and ω is a complex structure. To define the measure and the volume form we can use either the complex structure ω or the Riemannian metric η .

3.2.3 Average operator acting on sections

The averaging operation can be extended to a family of operators acting on sections of tensor bundles. This is especially important for the next *section*. Let $\pi^*, \pi_1, \pi_2, \pi^* \mathbf{T}^{(p,q)} \mathbf{M}$ and $\mathbf{T}^{(p,q)} \mathbf{M}$ be as before. Then let us consider the sections $S \in \Gamma(\mathbf{T}^{(p,q)} \mathbf{M})$ and $\pi^* S \in \Gamma(\pi^* \mathbf{T}^{(p,q)} \mathbf{M})$ and the isomorphisms

$$\pi_2|_v : \Gamma(\pi^* \mathbf{T}^{(p,q)} \mathbf{M}) \longrightarrow \Gamma(\mathbf{T}^{(p,q)} \mathbf{M}), \quad S_v \mapsto S_z,$$

$$\pi^* : \Gamma(\mathbf{T}^{(p,q)} \mathbf{M}) \longrightarrow \Gamma(\pi^* \mathbf{T}^{(p,q)} \mathbf{M}), \quad S_z \mapsto \pi_v^* S_z.$$

Both isomorphisms are defined pointwise.

Definition 3.2.2 Consider the family of fiber preserving endomorphisms

$$\{A(\mathbf{W}) : \Gamma(\pi^*\mathbf{TM}) \longrightarrow \Gamma(\pi^*\mathbf{TM}), \mathbf{W} \in \pi^{-1}(\mathbf{U}), \mathbf{U} \in \mathbf{M}\}.$$

The averaged operator of this family is the map

$$\langle A \rangle : \Gamma(\mathbf{T_U M}) \longrightarrow \Gamma(\mathbf{T_U M})$$

such that at each point $x \in \mathbf{U}$ it is given by:

$$(\langle A \rangle \cdot S)(x) := \frac{1}{\text{vol}(\mathbf{N}_x)} \left(\int_{\mathbf{N}_x} \pi_2|_u (A\pi^* \cdot S)(u) \right),$$

$$u \in \pi^{-1}(x), S \in \Gamma\mathbf{TM},$$

where $(A\pi^* \cdot S)(u)$ is the evaluation of the section $A(\pi^* \cdot S)$ at u .

A similar definition holds if the operators act on cartesian products of $\Gamma(\pi^*\mathbf{T}^{(p,q)}\mathbf{M})$.

3.3 Averaged connection of a linear connection on $\pi^*\mathbf{TM}$

We adopt a differential volume form $f(x, y) \omega_x(y)$ such that $(d\omega_x(y))|_{\mathbf{N}_x} = 0$. Therefore we denote $\omega_x(y)|_{\mathbf{N}_x} = d\text{vol}(x, y)$. Let us assume that a non-linear connection is defined on Σ , with $\Sigma \longrightarrow \mathbf{M}$ a vector bundle.

Definition 3.3.1 Let \mathbf{M} be a n -dimensional smooth manifold, $\pi(u) = x$ and consider a differentiable real function $f \in \mathcal{F}(\mathbf{M})$. Then $\pi^*f \in \mathcal{F}(\Sigma)$ is defined by the condition

$$\pi_u^*f = f(x). \quad (3.3.1)$$

The horizontal lift of the tangent vector $X^i \frac{\partial}{\partial x^i}|_x \in \mathbf{T}_x\mathbf{M}$ is

$$h : \Gamma\mathbf{TM} \longrightarrow \Gamma\mathbf{TN} \quad (3.3.2)$$

$$X^i \frac{\partial}{\partial x^i}|_x \mapsto X^i \frac{\delta}{\delta x^i}|_u, \quad u \in \pi^{-1}(x). \quad (3.3.3)$$

Proposition 3.3.2 *Let \mathbf{M} be a n -dimensional manifold and assume that \mathbf{N} is endowed with a non-linear connection, $u \in \pi^{-1}(x) \subset \mathbf{N}$, with $x \in \mathbf{M}$. Let us consider a linear connection ∇ defined on the vector bundle $\pi^*\mathbf{TM} \rightarrow \mathbf{N}$. Then a linear covariant derivative along $X \in \mathbf{T}\mathbf{M}$ is defined on \mathbf{M} , and is determined by the following conditions:*

1. $\forall X \in \mathbf{T}_x\mathbf{M}$ and $Y \in \Gamma\mathbf{TM}$, the covariant derivative of Y in the direction X , is given by the following averaging operations:

$$\langle \nabla \rangle_X Y := \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* Y \rangle_u, \quad \forall v \in \mathbf{U}_u, \quad (3.3.4)$$

where \mathbf{U}_u is an open neighborhood of $u \in \pi^{-1}(x)$.

2. For every smooth function $f \in \mathcal{FM}$ the covariant derivative is given by the following average:

$$\langle \nabla \rangle_X f := \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* f \rangle_u, \quad \forall v \in \mathbf{U}_u. \quad (3.3.5)$$

Proof: it is shown in reference [22, section 4] or in the *appendix*. \square

For the physical examples that we are interested, the notion of volume that we use is obtained by isometric embedding of the ambient Lorentzian structure η on the unit hyperboloid times a positive weight function f . The function $f_x := f(x, \cdot)$ will be required later to be at least $L^1(\Sigma_x)$ and with compact support Σ_x . This implies that the volume function is finite,

$$vol(\Sigma_x) := \int_{\Sigma_x} f(x, y) dvol(x, y) < \infty.$$

The manifold \mathbf{N}_x is oriented. In particular, the integration is performed in the unit tangent hyperboloid, which is

$$\mathbf{N}_x := \{y \in \mathbf{T}_x\mathbf{M}, \mid \eta(y, y) = 1, y^0 > 0\}.$$

Note that *proposition (3.3.2)* also holds in the more general case where the function f is not bounded and does not have compact support, but all the relevant integrals (in particular the volume function and average of the connection coefficients) are finite.

Definition 3.3.3 (*Generalized Torsion*) Let ∇ be a linear connection on $\pi^*\mathbf{TM} \rightarrow \mathbf{N}$, then the generalized torsion tensor acting on the vector fields $X, Y \in \mathbf{TM}$ is defined as

$$Tor(\nabla) : \Gamma\pi^*\mathbf{TM} \times \Gamma\pi^*\mathbf{TM} \longrightarrow \Gamma\pi^*\mathbf{TM}$$

$$(\pi^*X, \pi^*Y) \longrightarrow Tor(\nabla)(\pi^*X, \pi^*Y) = \nabla_{h(X)}\pi^*Y - \nabla_{h(Y)}\pi^*X - \pi^*[X, Y]. \quad (3.3.6)$$

This tensor is similar to the usual torsion tensor Tor ,

$$Tor(\nabla) : \Gamma\mathbf{TM} \times \Gamma\mathbf{TM} \longrightarrow \Gamma\mathbf{TM}$$

$$(X, Y) \longrightarrow Tor(\nabla)(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]. \quad (3.3.7)$$

Proposition 3.3.4 The averaged connection $\langle \nabla \rangle$ has a torsion $Tor(\langle \nabla \rangle)$ such that

$$Tor(\langle \nabla \rangle) = \langle Tor(\nabla) \rangle. \quad (3.3.8)$$

Proof: It is shown in reference [22] and in the *appendix* of this thesis. \square

Corollary 3.3.5 Let \mathbf{M} be an n -dimensional manifold and ∇ a linear connection on the bundle $\pi^*\mathbf{TM} \rightarrow \mathbf{M}$ with $Tor(\nabla) = 0$. Then $Tor(\langle \nabla \rangle) = 0$.

Proof: It is directly shown with the proof of *proposition (3.3.4)*. \square

If $Tor(\langle \nabla \rangle) = 0$ we say that the connection $\langle \nabla \rangle$ is torsion free.

Corollary 3.3.6 Let \mathbf{M} be an n -dimensional manifold. If the connection ∇ on $\pi^*\mathbf{TM}$ has the connection coefficients $\Gamma^i_{jk}(x, y)$, then the averaged connection \langle

$\nabla >$ has the coefficients

$$< \Gamma^i_{jk} > (x) = \frac{1}{\text{vol}(\mathbf{N}_x)} \int_{\mathbf{N}_x} \Gamma^i_{jk}(x, y) d\text{vol}(x, y). \quad (3.3.9)$$

Proof: Let $\{e_i\}$, $\{\pi^*e_i\}$, $\{h(e_i)\}$ be local frames for the sections of the vector bundles \mathbf{TM} , $\pi^*\mathbf{TM}$ and the horizontal bundle \mathcal{H} respectively, such that the covariant derivative is defined through the relations:

$$\nabla_{h(e_j)} \pi^* e_k = \Gamma^i_{jk} \pi^* e_i, \quad i, j, k = 0, 1, 2, \dots, n-1.$$

Then let us take the covariant derivative

$$\begin{aligned} < \nabla >_{e_j} e_k &= \frac{1}{\text{vol}(\mathbf{N}_x)} \left(\int_{\mathbf{N}_x} \pi_2(\nabla_{\iota(e_j)} \pi^* e_k) d\text{vol}_x(y) \right) = \frac{1}{\text{vol}(\mathbf{N}_x)} \left(\int_{\mathbf{N}_x} \pi_2 \Gamma^i_{jk} \pi^* e_i d\text{vol}_x(y) \right) = \\ &= \frac{1}{\text{vol}(\mathbf{N}_x)} \left(\int_{\mathbf{N}_x} \Gamma^i_{jk} d\text{vol}_x(y) \right) e_i. \end{aligned}$$

The relation (3.3.8) follows from the definition of the connection coefficients of the averaged connection,

$$< \nabla >_{e_j} e_k(x) = < \Gamma >^i_{jk}(x) e_i(x)$$

□

Chapter 4

Comparison of the Lorentz force equation and the averaged Lorentz force equation

4.1 Introduction

We start this *chapter* considering several aspects of two different, although related topics. The first one is the notion of *(semi)-Randers space* in the category of metrics with indefinite signature. The second deals with a geometric interpretation of the Lorentz force equation and the associated *averaged Lorentz equation*. Both are related and the discussion of the first topic helps to understanding the second.

The discussion of the above leads to a framework where the results on the averaged Lorentz connection can be formulated properly. We also introduce a metric structure in the space of connections on some pull-back bundles. Using this metric structure, it is possible to compare geodesics of the Lorentz connection and averaged Lorentz connection, which is the main result of this *chapter*.

The notion of Randers space, which introduces a non-reversible space-time structure, can be traced back to the original work by G. Randers [26]. The main moti-

vation was the investigation of a geometric structure encoding the time asymmetry. One of the results of that study was a unifying theory of gravitation and electrodynamics of charged point particles, encoding both physical interactions in an unifying space-time metric structure.

However, one can consider that the non-degeneracy of the associated *metric tensor* $g(x, y)$ was not discussed in detail, in particular when the issue of the gauge invariance associated with the electromagnetic potential is also considered. From a physical point of view, gauge invariance is a natural requirement. The combination of this requirement with the non-degeneracy criterion is non-trivial. Combining both has lead us to define Randers spaces in the context of pre-sheaf theory (this relation between Randers spaces and pre-sheave theory is only slightly treated in this thesis, since the details are still under construction). We hope that the identification of the appropriate formalism could provide a tool to formulate problems on Randers spaces in a consistent way.

There are other difficulties associated with the signature of the tensor $g(x, y)$. Indeed, while for positive definite Finsler metrics there is a satisfactory treatment (for instance [24, *chapter 11*]), for indefinite signatures, the theory of Finsler spaces is less universally accepted and several proposals are currently being used in the literature.

There are two general formalisms for indefinite Finsler spaces (that we call semi-Finsler structures): Asanov's formalism [27] and Beem's formalism [28], [29]. We will argue why both treatments and the corresponding physical interpretations are unsatisfactory, in particular when we try to apply them to Randers-type spaces. We can see the main problem with Asanov's definition when one considers gauge invariance issues related to the structure of Randers-type metrics. Also when one considers the possibility of light-like geodesics. The major problem with Beem's formalism is that there is no natural definition of Randers-type metric in that formalism. This is because Beem's formalism is based on homogeneous functions of degree two in the velocity variables y , while Randers-type functions are by definition

homogeneous of degree one in y .

In *section 4.3* we will provide a definition of semi-Randers space which is gauge invariant [14]. It has the advantage that all the notions involved are obtained directly from the Lorentz force equation and that it is a gauge invariant definition. However, this definition does not correspond to a Finsler or Lagrange structure. Indeed, we show that even being possible, there are severe practical difficulties to find a Lagrangian definition of semi-Randers space which is at the same time gauge invariant and globally defined in the tangent space \mathbf{TM} . This happens even in the absence of topological obstructions like the existence of *monopoles* for the 1-form A . Due to these difficulties we adopted a non-Lagrangian point of view in defining semi-Randers spaces.

In *section 4.4* we will propose a geometric description of the dynamics of one charged point particle interacting with an external electromagnetic field. This interpretation is natural in the framework for Randers-type space discussed in *section 4.3*. All relevant geometric data is extracted from the semi-Riemannian metric η and from the Lorentz force equation, which in an arbitrary local coordinate system reads

$$\frac{d^2\sigma^i}{d\tau^2} + {}^\eta\Gamma^i{}_{jk} \frac{d\sigma^j}{d\tau} \frac{d\sigma^k}{d\tau} + \eta^{ij}(dA)_{jk} \frac{d\sigma^k}{d\tau} \sqrt{\eta\left(\frac{d\sigma}{d\tau}, \frac{d\sigma}{d\tau}\right)} = 0, \quad i, j, k = 0, 1, 2, 3, \quad (4.1.1)$$

where $\sigma : \mathbf{I} \rightarrow \mathbf{M}$ is a solution curve on \mathbf{M} , ${}^\eta\Gamma^i{}_{jk}$ are the coefficients of the Levi-Civita connection ${}^\eta\nabla$ of η , dA is the exterior derivative of the 1-form A and the parameter τ is the proper-time of η along the curve σ . Then we interpret the equations (4.1.1) as the auto-parallel condition of a linear connection (in a convenient bundle). We called it the Lorentz connection. This interpretation does not make any additional assumption beyond the information already contained in the system (4.1.1) and the space-time metric η , except for some additional constraints on the generalized torsion, necessary to determine the connection coefficients completely.

In *section 4.5* we obtain the averaged Lorentz connection associated with the Lorentz connection, applying the averaging method discussed in *chapter 3*.

In *section 4.6* we will compare the solutions of the original system (4.1.1) and those of the auto-parallel curves of the averaged Lorentz connection. The result is that for the same initial conditions, in the *ultra-relativistic limit* and for narrow *one-particle* probability distribution functions, the solutions of both differential equations remain similar, even after a long time evolution, since there is a competition between time evolution and other factors. Therefore the original Lorentz force equation can be approximated by the averaged Lorentz force equation.

Remark. The natural object extracted from equation (4.1.1) is what we call *almost connection* (see the *appendix* for a formalization of the notion). In spite of this subtlety, we use through almost all the thesis the name connection (strictly speaking projective connection), since most of the calculations that we perform are also suitable for almost-connections (or *projective almost-connections*).

4.1.1 On the physical interpretation of the formalism

In the next sections we present a formal theory. However, the way the results are constructed depends on the physical problems which motivated them. The main problem was to model the dynamics of a bunch of particles in an accelerator machine, under the action of an external electromagnetic field. In this *chapter* we will consider the point particle dynamics point of view, which is related with the system of differential equations (4.1.1).

There are some hypotheses in the results that we consider which are related with the main problem, although it is not explicitly mentioned:

1. Bound conditions. For instance, all the parameter and variables that will appear in our results are assumed bounded in compact domains \mathbf{K} of the space-time \mathbf{M} . The reason for this assumption is that we are trying to model systems like a bunch of particles in accelerator machines. The evolution of a bunch starts at a given instant such that $t = 0$ with the *injection* and *separation* process of different bunches and the final time $t = T$, where the bunch reaches the target. Every coordinate and parameter of the model is bounded in $\mathbf{K} \subset \mathbf{M}$. The external electromagnetic fields are also bounded.

2. Ultra-relativistic regime. This is because the type of systems that we are describing are ultra-relativistic. We will define an energy function $E(x)$ which resembles the energy function used in accelerator physics. The ultra-relativistic regime happens when $E(x(t)) \gg 1$, where the mass of the specie of particle composing the bunch is set equal to 1 and $c = 1$.
3. Narrow distributions. This is one of the characteristics of the bunches in an accelerator machine. The narrowness of the distribution function is defined through the diameter α of the distribution function in the velocity space. The narrowness condition means that this diameter is small compared to the rest mass of the charged particles, $\alpha \ll 1$.
4. Adiabatic evolution. It is true that the change in energy is very slow compared to the energy itself, once the ultra-relativistic regimen has been reached. This is expressed by the condition $\frac{d \log E}{dt} \ll 1$.

The following reasons show why we have adopted the system of differential equations (4.1.1) as starting point for our geometrization of the electrodynamics of point particles are:

1. It seems that there is not a satisfactory and simple geometrization metric formalism for the interaction of a charged particle with an interacting external electromagnetic field (this is the main conclusion of sections 4.1-4.3).
2. The system of differential equations (4.1.1) is simple, contains all the symmetries that we are interested in and describes all the phenomenology of the dynamics of the charged point classical particles.
3. There exists an standard theory of geometric differential equations and its associated non-linear connections.
4. This geometric theory of differential equations provides the framework to apply the geometric averaging procedure described in *chapter 3*.

4.2 Criticism of the notion of semi-Randers space as space-time structure

4.2.1 Randers spaces as space-time structures

Before moving to the more specific problem of defining semi-Randers spaces, let us discuss the notion of semi-Finsler structure. Let \mathbf{M} be a \mathcal{C}^∞ n -dimensional manifold, \mathbf{TM} its tangent bundle manifold with $\mathbf{TM} \supset \mathbf{N}$ and with projection $\pi : \mathbf{N} \longrightarrow \mathbf{M}$, the restriction to \mathbf{N} of the canonical projection $\pi : \mathbf{TM} \longrightarrow \mathbf{M}$. Therefore, (\mathbf{N}, π) is a sub-bundle of \mathbf{TM} .

Let us consider the following two standard definitions of semi-Finsler structures currently being used in the literature:

1. *Asanov's definition* [27],

Definition 4.2.1 *A semi-Finsler structure F defined on the n -dimensional manifold \mathbf{M} is a positive, real function $F : \mathbf{N} \longrightarrow]0, \infty[$ such that:*

- (a) *It is smooth in \mathbf{N} ,*
- (b) *It is positive homogeneous of degree 1 in y , $F(x, \lambda y) = \lambda F(x, y)$, $\forall \lambda > 0$,*
- (c) *The vertical Hessian matrix*

$$g_{ij}(x, y) := \frac{1}{2} \frac{\partial^2 F^2(x, y)}{\partial y^i \partial y^j} \quad (4.2.1)$$

is non-degenerate on \mathbf{N} .

$g_{ij}(x, y)$ is the matrix of the fundamental tensor. The set \mathbf{N}_x is the admissible set of tangent vectors at x ; the disjoint union $\mathbf{N} = \bigsqcup_{x \in \mathbf{M}} \mathbf{N}_x$ is the admissible set of vectors over \mathbf{M} .

In the particular case when the manifold is 4-dimensional and (g_{ij}) has signature $(+, -, -, -)$, the pair (\mathbf{M}, F) is a Finslerian space-time.

2. *Beem's definition* [28], [29]

Definition 4.2.2 *A semi-Finsler structure defined on the n -dimensional manifold \mathbf{M} is a real function $L : \mathbf{TM} \longrightarrow \mathbf{R}$ such that*

- (a) *It is smooth in the slit tangent bundle $\tilde{\mathbf{N}} := \mathbf{TM} \setminus \{0\}$*
- (b) *It is positive homogeneous of degree 2 in y , $L(x, \lambda y) = \lambda^2 L(x, y)$, $\forall \lambda > 0$,*
- (c) *The Hessian matrix*

$$g_{ij}(x, y) := \frac{1}{2} \frac{\partial^2 L(x, y)}{\partial y^i \partial y^j} \quad (4.2.2)$$

is non-degenerate on $\tilde{\mathbf{N}}$.

In the particular case when the manifold is 4-dimensional and g_{ij} has signature $(+, -, -, -)$, the pair (\mathbf{M}, L) is a Finslerian space-time.

4.2.2 Comparison of Asanov's and Beem's formalism

Some differences between the above definitions are highlighted below:

1. In Beem's framework there is a geometric definition of light-like vectors and it is possible to derive Finslerian geodesics, including light-like geodesics, from a variational principle [30]. By construction, in Asanov's formalism it is not possible to do that in an invariant way, because light-like vectors are excluded in the formalism from the beginning, since nothing is said about how to extend the function F^2 from \mathbf{N} to \mathbf{TM} .
2. Let $\Theta(\mathbf{M})$ be the set of all piecewise smooth paths $\sigma : \mathbf{I} \longrightarrow \mathbf{M}$. In Asanov's framework, given a parameterized path $\sigma : \mathbf{I} \longrightarrow \mathbf{M}$, $\mathbf{I} \subset \mathbf{R}$ on the semi-Finsler manifold (\mathbf{M}, F) such that $\dot{\sigma} \in \mathbf{N}$ for all $t \in \mathbf{I}$, the length functional acting on σ is given by the following expression:

$$\mathcal{E}_A : \Theta(\mathbf{M}) \longrightarrow \mathbf{R}$$

$$\sigma(t) \mapsto \mathcal{E}_A(\sigma) := \int_{t_{min}}^{t_{max}} F(\sigma(t), \dot{\sigma}(t)) dt, \quad \mathbf{I} = [t_{min}, t_{max}]. \quad (4.2.3)$$

Due to the homogeneous condition of the Finsler function F , \mathcal{E}_A is a re-parametrization invariant functional. On the other hand, if we consider Beem's

definition, the energy functional is given by the following expression [30]:

$$\mathcal{E}_B : \Theta(\mathbf{M}) \longrightarrow \mathbf{R}$$

$$\sigma(t) \longrightarrow \mathcal{E}_B(\sigma) := \int_{t_{min}}^{t_{max}} L(\sigma(t), \dot{\sigma}(t)) dt, \quad \mathbf{I} = [t_{min}, t_{max}]. \quad (4.2.4)$$

Formulated in this way, Beem's energy functional is not re-parametrization invariant, because the fundamental function L is homogeneous of degree two in y .

3. A third difference emerges when we consider the category of Randers-type spaces:

Definition 4.2.3 (*semi-Randers Space as semi-Finsler Space*)

In Asanov's framework, a semi-Randers space is characterized by a semi-Finsler function of the form:

$$F(x, y) = \sqrt{\eta_{ij}(x)y^i y^j} + A_i(x)y^i, \quad (4.2.5)$$

where $\eta_{ij}(x)dx^i \otimes dx^j$ is a semi-Riemannian metric defined on \mathbf{M} and $A(x, y) := A_i(x)y^i$ is the result of the action of the 1-form $A(x) = A_i(x)dx^i$ on $y \in \mathbf{TM}$.

In the positive definite case and when η is a Riemannian metric, the requirement that g_{ij} is non-degenerate implies that the 1-form (A_1, \dots, A_n) is bounded by η :

$$A_i A_j \eta^{ij} < 1, \quad \eta^{ik} \eta_{kj} = \delta^i_j.$$

The indefinite case is quite different, since there is not a natural Riemannian metric that induces a norm in the space of homomorphisms. Therefore, the criterion for non-degeneracy is not trivial and further structure is required.

Secondly, for both the positive definite and indefinite metric, only the variation of the length functional (4.2.3) (for fixed initial and final point variations) is invariant under the gauge transformation $A \mapsto A + d\lambda$ (not directly the integrand itself); the Finsler function (4.2.5) is not gauge invariant as well.

Even in the case that we could define a metric and norm, transforming the 1-form A by a gauge transformation can change the norm and therefore the hessian (g_{ij}) can become degenerate.

On the other hand, the notion of Randers space in Beem's formalism is even more problematic. In this case there is not a formulation of semi-Randers spaces (because eq. (4.2.5) is positive homogeneous of degree one in y). This suggests that a proper formulation of the notion of semi-Randers space requires going beyond metric structures.

4. It is interesting to have a definition of semi-Randers structure capable of taking light-like trajectories for charged particles into account. As we have seen, Asanov's treatment is not able to consider light-like vectors. On the other hand, Beem's formalism is not capable of considering this problem, since there is not a known Randers-type structure in Beem's formalism. However, the asymptotic expansion of the ultra-relativistic charged cold fluid model presented in [6] is an example where those light-like trajectories appear naturally. In that model, the leading order contribution to the mean velocity field of the charged cold fluid is a light-like velocity vector field, interacting with the external electromagnetic field; perturbative corrections change the velocity vector field to a time-like vector field.

The above observations make it reasonable to introduce a non-metric interpretation for semi-Randers spaces. The option that we have adopted has been to formalize a geometric structure from the geometric and physical data that we have: the Lorentz force equation and the Lorentzian metric η . This will lead us to solve some of the problems mentioned before. It provides a rigorous framework to discuss further developments.

4.3 Non-Lagrangian notion of semi-Randers space

4.3.1 Non-Lagrangian notion of semi-Randers space

Let us assume the existence of a smooth semi-Riemannian structure η on the manifold \mathbf{M} . This implies that the function

$$\eta : \mathbf{TM} \times \mathbf{TM} \longrightarrow \mathbf{R}$$

$$(X, Y) \mapsto \eta_{ij}(x) X^i Y^j, \quad X, Y \in \mathbf{T}_x \mathbf{M}$$

is smooth in the variables x, X^i, Y^j . Since we will use the square root $\sqrt{\eta_{ij}(x) X^i Y^j}$, we also require that $\sqrt{\eta_{ij}(x) X^i X^j}$ is smooth in ${}^\eta \mathbf{N} := \bigcup_{x \in \mathbf{TM}} \{X \in \mathbf{T}_x \mathbf{M}, \eta_{ij}(x) X^i X^j > 0\}$. The null-cone is ${}^\eta \mathbf{NC} := \bigsqcup_{x \in \mathbf{M}} \{y \in \mathbf{T}_x \mathbf{M} \mid \eta(y, y) = 0\}$. We propose a notion of semi-Randers space based on the following

Definition 4.3.1 *A semi-Randers space consists of a triplet $(\mathbf{M}, \eta, \mathbf{F})$, where \mathbf{M} is a space-time manifold, η is a semi-Riemannian metric continuous on \mathbf{TM} and smooth on $\mathbf{TM} \setminus {}^\eta \mathbf{NC}$ and a 2-form $\mathbf{F} \in \bigwedge^2 \mathbf{M}$ such that $d\mathbf{F} = 0$.*

\mathbf{F} is in the second de Rham cohomology group $H^2(\mathbf{M})$. Due to Poincaré's lemma, there is a locally smooth 1-form A such that $dA = \mathbf{F}$. Any pair of locally smooth 1-forms \tilde{A} and A such that $\tilde{A} = A + d\lambda$, with λ a locally smooth real function defined on the given open neighborhood, are *equivalent* and produce under exterior derivative the same cohomology class $[\mathbf{F}] \in H^2(\mathbf{M})$ that contains the element \mathbf{F} : $d(d\lambda + A) = dA = \mathbf{F}$. Note that we are speaking of locally smooth 1-forms A and of globally smooth 2-forms \mathbf{F} . Therefore, instead of giving \mathbf{F} , one can consider the equivalence class of 1-forms A ,

$$[A] := \{\tilde{A} = A + d\lambda, dA = \mathbf{F} \text{ in the intersection of the opens sets where } A \text{ and } \lambda \text{ are defined}\},$$

with A a locally smooth 1-form defined on the open set $\mathbf{U} \subset \mathbf{M}$, \tilde{A} a locally smooth 1-form defined on the open set $\tilde{\mathbf{U}} \subset \mathbf{M}$ and λ a locally smooth function defined on $\mathbf{U} \cap \tilde{\mathbf{U}}$. Then if two locally smooth forms 1-forms ${}^\mu A$ and ${}^\nu A$, representatives of $[A]$,

are defined on ${}^\mu\mathbf{U}$ and ${}^\nu\mathbf{U}$ respectively, one has that $d({}^\mu A - {}^\nu A) = 0$. For each point of the open neighborhood ${}^{\mu\nu}\mathbf{U} = {}^\mu\mathbf{U} \cap {}^\nu\mathbf{U}$, there is a locally smooth function defined in an open neighborhood of $x \in \mathbf{U}(x) \subset {}^{\mu\nu}\mathbf{U}$ such that $({}^\mu A - {}^\nu A) = d({}^{\mu\nu}\lambda)$ (a consequence of the Poincaré lemma).

Based on these arguments, we give an alternative definition of semi-Randers space:

Definition 4.3.2 *A semi-Randers space consists of a triplet $(\mathbf{M}, \eta, [A])$, where \mathbf{M} is a space-time manifold, η is a semi-Riemannian metric continuous on \mathbf{TM} and smooth on $\mathbf{TM} \setminus {}^\eta\mathbf{NC}$ and the class of locally smooth 1-forms A is defined such that $dA = \mathbf{F}$ for any $A \in [A]$.*

Proposition 4.3.3 *These definitions of semi-Randers space are equivalent.*

Proof. We proved already one of the directions of the equivalence. To show the other direction, one needs to construct locally 1-forms which produce the required 2-form \mathbf{F} under exterior differentiation. This is achieved by the Poincaré lemma in a star-shaped domain [32, pg 155-156]. The formula for the 1-form A is

$$A(x) = \left(\int_0^1 t \sum_{k=0}^{n-1} x^k \mathbf{F}_{kj}(tx) dt \right) dx^j.$$

□

We will adopt *definition 4.3.2*, since it has the advantage that it allows a discussion of some local issues related with the inverse variational problem of the Lorentz force equation. Essentially, this is the reason that even if the topology of \mathbf{M} is trivial, the 1-forms A are only locally smooth.

As we have learned from the discussion above, a proper treatment of semi-Randers spaces combined with gauge invariance requires consideration of locally smooth potentials. There are also locally smooth functions and local *compatibility conditions*. These kind of structures are formalized by the notion of pre-sheaf structure (and the related notion of sheaf structure) [32, 33].

Let us denote the set of locally smooth functions over \mathbf{M} by $\bigwedge_{loc}^p \mathbf{M}$. This is a pre-sheaf structure. The pre-sheaf of locally smooth functions on open sets of \mathbf{M} is denoted by $\mathcal{F}_{loc}(\mathbf{M})$. Given a Lorentz semi-Randers structure $(\mathbf{M}, \eta, [A])$, for each of the representatives $A \in [A] \in \bigwedge_{loc}^1 \mathbf{M}$, there is on \mathbf{M} a function F_A defined by the following expression:

$$F_A(x, y) = \begin{cases} \sqrt{\eta_{ij}(x)y^i y^j} + A_i(x)y^i & \text{for } \eta_{ij}(x)y^i y^j \geq 0, \\ \sqrt{-\eta_{ij}(x)y^i y^j} + A_i(x)y^i & \text{for } \eta_{ij}(x)y^i y^j \leq 0. \end{cases} \quad (4.3.1)$$

The following properties follow easily from definition (4.3.1) and from the definition F_A :

Proposition 4.3.4 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space with η a semi-Riemannian metric, $A \in [A]$ and F_A given by equation (4.3.1). Then*

1. *On the null cone $\mathbf{NC}_x := \{y \in \mathbf{T}_x \mathbf{M} \mid \eta_{ij}(x)y^i y^j = 0\}$ F_A is of class \mathcal{C}^0 , for $\eta \in \mathcal{C}^0$ and $A \in \bigwedge_{loc}^p \mathbf{M}$.*
2. *The subset where $\eta_x(y, y) \neq 0$ is an open subset of $\mathbf{T}_x \mathbf{M}$ and F_A is smooth on $\mathbf{T}_x \mathbf{M} \setminus \mathbf{NC}_x$, for η smooth and $A \in \bigwedge_{loc}^1 \mathbf{M}$.*
3. *The function F_A is positive homogeneous of degree 1 in y .*

Remark. There is no constraint on the non-degeneracy of the fundamental tensor g_{ij} . Therefore, it is not required that any representative A of $[A]$ be bounded by 1, as of is the case for positive definite Randers spaces [24, chapter 11].

4.3.2 Variational principle on semi-Randers spaces

Let ${}^t\Theta(\mathbf{M})$ be the set of piecewise smooth curves on \mathbf{M} with time-like tangent vector field. The functional acting on σ is given by the integral:

$$\mathcal{E}_{F_A} : \Theta(\mathbf{M}) \longrightarrow \mathbf{R}$$

$$\sigma \mapsto \mathcal{E}_{F_A}(\sigma) := \int_{\sigma} F_A(\sigma(\tau), \dot{\sigma}(\tau)) d\tau \quad (4.3.2)$$

with τ the proper time associated with η along the curve σ . The functional is gauge invariant up to a constant: if we choose another representative $\tilde{A} = A + d\lambda$, then $\mathcal{E}_{F_A}(\sigma) = \mathcal{E}_{F_{\tilde{A}}}(\sigma) + \text{constant}$, the constant coming from the boundary terms of the integral. Therefore, the variation of the functional is well defined on a given semi-Randers space $(\mathbf{M}, \eta, [A])$, for fixed initial and final points variations.

In order to guarantee the construction of the first variation formula and the existence and uniqueness of the corresponding solution it is necessary that the vertical Hessian g_{ij} be non-degenerate [31]. However, given a representative $A \in [A]$, one can not guarantee that the Hessian of F_A is non-degenerate. Due to the possibility of doing gauge transformations in the representative $A(x) \mapsto A(x) + d\lambda(x)$ we have

Proposition 4.3.5 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space. Assume that the image of the curve σ on the manifold \mathbf{M} is a compact subset. Then*

1. *There is a representative $\bar{A} \in [A]$ such that the Hessian of the functional $F_{\bar{A}}$ is non-degenerate.*
2. *The functional (4.3.2) is well defined on the Randers space $(\mathbf{M}, \eta, [A])$, except for a constant depending on the representative $\bar{A} \in [A]$.*
3. *If the geodesic curves are parameterized by the proper time associated with the Lorentzian metric η , the Euler-Lagrange equation of the functional F_A is the Lorentz force equation.*

Proof: There are several steps in the proof:

1. Using the gauge invariance of $\mathcal{E}_F(\sigma)$ up to a constant, we can obtain locally an element $\bar{A} \in [A]$ such that $|\bar{A}_i \bar{A}_j \eta^{ij}| < 1$ in an open neighborhood in the following way. Consider that we start with a 1-form A which is not bounded by 1. The 1-form $\bar{A}(x) = A(x) + d\lambda(x)$ is also a representative of $[A]$. The requirement that the hessian of \bar{A} is non-degenerate is, using a generalization of the condition [24, pg 289]

$$0 < \left| 2 + \frac{\bar{A}_i(x)y^i + \sqrt{\eta_{ij}(x)y^i y^j} (\eta^{ij} \bar{A}_i(x) \bar{A}_j(x))}{\sqrt{\eta_{ij}(x)y^i y^j} + \bar{A}_i(x)y^i} \right|$$

This condition is obtained in [24] relating the determinant of the metric η_{ij} and the determinant of g_{ij} , the fundamental tensor of a Randers-type metric structure. Although the authors are considering positive definite metrics, the result is valid for arbitrary signatures of the metric η_{ij} , for $y \in \mathbf{N}_x$

On the unit tangent hyperboloid Σ_x this condition reads

$$0 < \left| 2 + \frac{\bar{A}_i(x)y^i + \eta^{ij}(x)\bar{A}_i(x)\bar{A}_j(x)}{1 + \bar{A}_i(x)y^i} \right|.$$

Let us assume (if it is negative, the treatment is similar) that

$$\epsilon^2(x, y) := 2 + \frac{\bar{A}_i(x)y^i + \eta^{ij}(x)\bar{A}_i(x)\bar{A}_j(x)}{1 + \bar{A}_i(x)y^i} > 0.$$

To write down this condition, one needs that $1 + \bar{A}_i(x)y^i \neq 0$; the region where this does not hold is the intersection of the hyperplane

$$\mathbf{P}_x := \{y \in \mathbf{T}_x\mathbf{M} \mid 1 + \bar{A}_i(x)y^i = 0\}$$

with the unit hyperboloid Σ_x . The intersection is such that $\mathbf{P}_x \cap \Sigma_x \subset \{y \in \Sigma_x \mid F(x, y) = 0\}$.

Let us write in detail the above condition of positiveness for the potential $\bar{A}_i = A_i(x) + \partial_i\lambda(x)$:

$$\epsilon^2(x, y) = 2 + \frac{(A_i(x) + \partial_i\lambda(x))y^i + \eta^{ij}(x)(A_i(x) + \partial_i\lambda(x))(A_j(x) + \partial_j\lambda(x))}{1 + (A_i(x) + \partial_i\lambda(x))y^i}.$$

If $y \in \Sigma_x$, then for $\beta \neq 1$, βy is not in Σ_x . The situation is different if $y \in \mathbf{NC}_x$. Then $\beta y \in \mathbf{NC}_x$, even if $\beta \neq 1$. Now we make the approximation $\Sigma_x \longrightarrow \mathbf{NC}_x$ in the asymptotic limit $y^0 \longrightarrow \infty$. In this approximation, one can perform limits in the expression for $\epsilon^2(x, y)$. In particular, one can consider $\epsilon^2(x, \beta y)$ for large β :

$$\lim_{\beta \rightarrow \infty} \epsilon^2(x, \beta y) = 2 + \lim_{\beta \rightarrow \infty} \frac{(A_i(x) + \partial_i\lambda(x))\beta y^i + \eta^{ij}(x)(A_i(x) + \partial_i\lambda(x))(A_j(x) + \partial_j\lambda(x))}{1 + (A_i(x) + \partial_i\lambda(x))\beta y^i} =$$

$$= 2 + \lim_{\beta \rightarrow \infty} \frac{(A_i(x) + \partial_i \lambda(x)) \beta y^i}{1 + (A_i(x) + \partial_i \lambda(x)) \beta y^i} = 3.$$

That this limit is well defined for large enough y^0 can be seen in the following way. Let us assume that $1 + (A_i(x) + \partial_i \lambda(x)) y^i = 0$. Then we consider the expression $1 + (A_i(x) + \partial_i \lambda(x)) \beta y^i$ for $\beta \gg 1$. It is impossible that the second expression is zero except if $\beta = 1$ and $1 + (A_i(x) + \partial_i \lambda(x)) y^i = 0$.

One should prove that $\epsilon^2(x, y)$ is positive for any $y \in \Sigma_x$. This is achieved because $\epsilon^2(x, y)$ is invariant. Therefore, we can change to a coordinate system where y^0 is arbitrary and the value of the bound does not change.

2. If the curve σ is such that its image $\sigma(\mathbf{I}) \subset \mathbf{M}$ is covered by several open sets, for instance ${}^\mu \mathbf{U}$ and ${}^\nu \mathbf{U}$, $\mathcal{F}_A(\sigma)$ is evaluated using both representatives. In principle, if we consider two representatives ${}^\mu \bar{A}$ and ${}^\nu \bar{A}$, there is a contribution coming from *boundary terms* coming from the evaluation of $d\lambda$ on those points in the intersection ${}^\mu \mathbf{U} \cap {}^\nu \mathbf{U}$. This contribution do not appear in the first variation of the functional. Therefore, the first variation of (4.3.2) exists and does not depend on the representative A , for fixed initial and final point variations of σ . The corresponding extremal curves exist and they are unique. They correspond to the solutions of the Lorentz force equation.
3. If the image $\sigma(\mathbf{I}) \subset \mathbf{M}$ is a compact set, the number of sets ${}^\mu \mathbf{U}$ that we need is finite. Then we obtain a globally defined section of $\bigwedge_{loc}^1 \mathbf{M}$.
4. Once a locally differentiable 1-form A with the required properties is obtained globally over the variation $Var(\sigma)$ of $\sigma : \mathbf{I} \longrightarrow \mathbf{M}$, one can follow the standard proof of the deduction of the Lorentz force equation from the variation of a functional [26], [2, pg 47-52]. The variation $Var(\sigma)$ must be constructed such that the 1-form A is globally defined on $Var(\sigma)$. \square

Remarks

1. It is important to notice that the non-degeneracy of the metric g_{ij} does not necessarily imply that it has the same signature as the semi-Riemannian metric

η . Further investigations are required to determine the criteria for conservation of signature.

2. If the curve σ is parameterized with respect to a parameter such that the Finslerian arc-length $F(\sigma, \dot{\sigma})$ is constant along the geodesic, the geodesic equations have a complicated form (see for instance [24, pg. 296]) and they are not invariant under arbitrary gauge transformations of the 1-form $A \longrightarrow A + d\lambda$, $\lambda \in \mathcal{F}_{loc}(\mathbf{M})$.
3. As we have mentioned before, the fact that we are speaking of local data natural leads us to consider notions from pre-sheaf and sheaf theory [32, *chapter 6*, 33, *chapter 2*] as the basic ingredient in the definition of Randers spaces. Sheaf theory is a theory that allows us to treat local objects, for example germs of locally smooth functions or locally smooth differential forms [33, *chapter 2*]. Hence, we think this is a natural framework to study the geometry and variational properties of semi-Randers spaces.

With definition (4.3.2) at hand, the problem of how to introduce the gauge symmetry in a Randers geometry is solved. The price to pay is

1. The class $[A]$ and the Riemannian metric η are unrelated geometric objects. This is in contradiction with the spirit of Randers spaces as a space-time asymmetric structure.
2. One needs to consider locally smooth 1-forms A instead of globally defined 1-forms. This happens even if the topology of \mathbf{M} is trivial or the cohomology class of A is trivial.

We have seen that given a 1-form $A \in [A]$ one has to *work hard* to find another representative $\bar{A} \in [A]$ such that Douglas's theorem [31] holds. Douglas's theorem states the condition under which a system of ordinary differential equations can be interpreted as the Euler-Lagrange equation coming from a Lagrangian. One of the requirements is that the vertical hessian of the Lagrangian must be non-degenerate, $\det(g_{ij}) \neq 0$.

On the other hand, light-like trajectories cannot be considered in this formalism for semi-Randers spaces in a natural way. Beem's formalism allows the treatment of light-like geodesics as extremal curves of an energy functional for some examples of indefinite Finsler space-times. However, it is not known if a semi-Randers function exists in Beem's formalism.

In conclusion, the advantage of the definition given here over Asanov's definition of semi-Randers space is that it is consistent with gauge invariance. Nevertheless, our formalism is still not completely satisfactory, since it can not be used for light-like trajectories.

4.4 Geometric formulation of the Lorentz force equation

4.4.1 The non-linear connection associated with the Lorentz force equation

Let us consider a semi-Randers space $(\mathbf{M}, \eta, [A])$ and the bundle $\mathbf{N} \longrightarrow \mathbf{M}$, $\mathbf{N} := \bigsqcup_{x \in \mathbf{M}} \{y \in \mathbf{T}_x \mathbf{M}, \eta(y, y) \geq 0\} \subset \mathbf{TM}$. Then the following diagram

$$\begin{array}{ccc} & & \mathbf{TM} \\ & \nearrow e & \downarrow \pi \\ \mathbf{N} & \xrightarrow{\pi} & \mathbf{M} \end{array}$$

where e is the following natural embedding

$$e : \mathbf{N} \longrightarrow \mathbf{TM}$$

$$(x, y) \mapsto (x, y), \quad x \in \mathbf{M}, \quad y \in \mathbf{N}_x$$

We will consider the differential map $d\pi : \mathbf{TN} \longrightarrow \mathbf{TM}$. Recall that the vertical bundle is defined as the kernel $\mathcal{V} := \ker(d\pi)$; at each point one has $\ker(d\pi|_u) := \mathcal{V}_u$, with $u \in \mathbf{N}$.

The system of second order differential equations (4.1.1) determines a special type of vector field on \mathbf{N} called spray. It is well known that a spray defines an Ehresmann connection on \mathbf{TN} [34]. Let us denote by $\eta(Z, Y) := \eta_{ij}(x)Z^i Y^j$.

Definition 4.4.1 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space. For each tangent vector $y \in \mathbf{T}_x \mathbf{M}$ with $\eta(y, y) > 0$, the following functions are well defined,*

$$\begin{aligned} {}^L\Gamma^i_{jk}(x, y) = & {}^\eta\Gamma^i_{jk}(x) + \frac{1}{2\sqrt{\eta(y, y)}}(\mathbf{F}^i_j(x)y^m\eta_{mk} + \mathbf{F}^i_k(x)y^m\eta_{mj}) + \\ & + \mathbf{F}^i_m(x)\frac{y^m}{2\sqrt{\eta(y, y)}}\left(\eta_{jk} - \frac{1}{\eta(y, y)}\eta_{js}\eta_{kl}y^s y^l\right), \end{aligned} \quad (4.4.1)$$

${}^\eta\Gamma^i_{jk}(x)$, $(i, j, k = 0, 1, 2, \dots, n)$ are the connection coefficients of the Levi-Civita connection ${}^\eta\nabla$ in a local frame, $\mathbf{F}_{ij} := \partial_i A_j - \partial_j A_i$ and $\mathbf{F}^i_j = \eta^{ik}\mathbf{F}_{kj}$, for any representative $A \in [A]$.

The unit hyperboloid sub-bundle Σ acquires an induced connection, whose connection coefficients are

$${}^L\Gamma^i_{jk}(x, y)|_\Sigma = {}^\eta\Gamma^i_{jk}(x) + \frac{1}{2}(\mathbf{F}^i_j(x)y^m\eta_{mk} + \mathbf{F}^i_k(x)y^m\eta_{mj}) + \mathbf{F}^i_m(x)\frac{y^m}{2}(\eta_{jk} - \eta_{js}\eta_{kl}y^s y^l).$$

The structure of these functions is clear: ${}^\eta\Gamma^i_{jk}(x)$ are the connection coefficients of the Lorentzian metric η ; the other two terms are tensorial. Indeed, one can define the following expressions:

$$\begin{aligned} L^i_{jk}(x, y) &= \frac{1}{2\sqrt{\eta(y, y)}}(\mathbf{F}^i_j(x)y^m\eta_{mk} + \mathbf{F}^i_k(x)y^m\eta_{mj}), \\ T^i_{jk}(x, y) &= \mathbf{F}^i_m(x)\frac{y^m}{2\sqrt{\eta(y, y)}}\left(\eta_{jk} - \frac{1}{\eta(y, y)}\eta_{js}\eta_{kl}y^s y^l\right). \end{aligned}$$

Therefore,

$${}^L\Gamma^i_{jk} = {}^\eta\Gamma^i_{jk} + T^i_{jk} + L^i_{jk}.$$

With the functions L^i_{jk} and T^i_{jk} , one can construct the following maps:

$$L_u : \mathbf{T}_u \mathbf{N} \times \mathbf{T}_u \mathbf{N} \longrightarrow \mathbf{T}_u \mathbf{N}$$

$$(X, Y) \mapsto L^i{}_{jk}(x, y) X^j Y^k \frac{\delta}{\delta x^i}.$$

Recall that the section $\frac{\delta}{\delta x^i}$ is the horizontal lift of the section ∂_i .

The second operator that we define is

$$T_u : \mathbf{T}_u \mathbf{N} \times \mathbf{T}_u \mathbf{N} \longrightarrow \mathbf{T}_u \mathbf{N}$$

$$(X, Y) \mapsto T^i{}_{jk}(x, y) X^j Y^k \frac{\delta}{\delta x^i}.$$

$u = (x, y)$ and X, Y are arbitrary tangent vectors $X, Y \in \mathbf{T}_u \mathbf{N}$. This can be generalized to homomorphisms acting on vector sections.

Note the following elementary property

$$T_u(Y, Y) = 0, \quad \forall y \in \mathbf{N}_x, \quad Y = y^i \frac{\delta}{\delta x^i}, \quad u = (x, y).$$

However, $T_u(\cdot, Y) \neq 0$ in general.

4.4.2 The Koszul connection ${}^L\nabla$ on \mathbf{TN} associated with the Lorentz force equation

Let $\{e_0, \dots, e_{n-1}\}$ be a local basis for the sections of the frame bundle associated with the tangent bundle $\Gamma \mathbf{TM}$ and let us assume that each e_i is a time-like tangent vector at the point $x \in \mathbf{M}$ (therefore, the metric cannot be diagonal in this basis, since η is a semi-Riemannian metric). Then $\{\pi^* e_0, \dots, \pi^* e_{n-1}\}$ is a local frame for the fiber $\pi_u^{-1} \subset \pi^* \mathbf{TM}$, $u \in \mathbf{N}$, $\{h_0, \dots, h_{n-1}\}$ is the local frame of the horizontal distribution $\mathcal{H}_u \subset \mathbf{T}_u \mathbf{N}$ obtained by the horizontal lift $h_i = h(e_i)$ and $\{v_0, \dots, v_{n-1}\}$ is a local frame for the vertical distribution $\mathcal{V}_u \subset \mathbf{T}_u \mathbf{N}$.

Given the set of functions $\{{}^L\Gamma^i{}_{jk}, i, j, k = 0, \dots, n-1\}$ and the non-linear connection associated with the system of differential equations, there is an associated Koszul connection on \mathbf{TN} [35] (a Koszul connection is a linear connection defined through the corresponding covariant derivative).

Proposition 4.4.2 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space and \mathbf{M} a n -dimensional*

manifold. There is defined a covariant derivative ${}^L D$ on \mathbf{TN} determined by the following conditions:

1. For each $X \in \mathcal{H}_u$ and $Z \in \Gamma\mathcal{H}$

$${}^L D_X Z = X^k {}^L \Gamma^i_{jk}(x, y) Z^j h_i, \quad X = X^i h_i|_u, \quad Z = Z^i h_i|_v,$$

with $\{h_i\}$ a local frame for the horizontal distribution, $u = (x, y)$ and v an arbitrary point of an open set $\tilde{\mathbf{O}} \subset \mathbf{N}$ containing u .

2. The covariant derivative of arbitrary sections $Z \in \Gamma\mathbf{TN}$ along vertical direction is zero:

$${}^L D_V Z = 0, \quad \forall V \in \mathcal{V}, Z \in \Gamma\mathbf{TN}.$$

3. The covariant derivative ${}^L D$ is symmetric, i.e, has zero horizontal torsion:

$${}^L D_U V - {}^L D_V U - [U, V] = 0, \quad \forall U, V \in \mathcal{H}.$$

4. For all $X \in \mathcal{H}_u$ and $Z \in \Gamma\mathcal{V}$, ${}^L D_X Z = 0$.

Proof: Let us consider the Finsler geodesic equation associated with the semi-Randers space $F_A = \sqrt{\eta_{ij}(x)y^i y^j} + A_i(x)y^i$ but parameterized using the arc-length of the Lorentzian metric η :

$$\frac{d^2 x^i}{d\tau^2} + \eta \Gamma^i_{jk} \frac{dx^j}{d\tau} \frac{dx^k}{d\tau} + \eta^{ij} (dA)_{jk} \sqrt{\eta\left(\frac{dx}{d\tau}, \frac{dx}{d\tau}\right)} \frac{dx^k}{d\tau} = 0,$$

where $\mathbf{F} = dA$ is the exterior differential of the 1-form A and $\eta(X, Z) = \eta_{ij}(x)X^i Z^j$. From these equations, we can read the value of the semi-spray coefficients, which are

$$G^i(x, y) = \eta \Gamma^i_{jk}(x) y^j y^k + \eta^{ij}(x) (dA)_{jk}(x) \sqrt{\eta(y, y)} y^k.$$

Taking the first and second derivatives with respect to y , we obtain

$$\frac{1}{2} \frac{\partial}{\partial y^j} G^i(x, y) = \eta \Gamma^i_{lj}(x) y^l + \eta^{il}(x) (dA)_{lm}(x) \frac{1}{\sqrt{\eta(y, y)}} \eta_{js}(x) y^s y^m + \eta^{il}(x) (dA)_{lj}(x) \sqrt{\eta(y, y)}.$$

$$\begin{aligned}
\frac{1}{2} \frac{\partial}{\partial y^k} \frac{\partial}{\partial y^j} G^i(x, y) &= \eta^i{}_{kj}(x) - \eta^{il}(x)(dA)_{lm}(x) \frac{1}{2(\eta(y, y))^{3/2}} \eta_{js}(x) y^s y^m \eta_{kp}(x) y^p + \\
&+ \eta^{il}(x)(dA)_{lm}(x) \frac{1}{2(\sqrt{\eta(y, y)})} \eta_{jk}(x) y^m + \eta^{il}(x)(dA)_{lk}(x) \frac{1}{2(\sqrt{\eta(y, y)})} \eta_{js} y^s + \\
&+ \eta^{il}(x)(dA)_{lj}(x) \frac{1}{2(\sqrt{\eta(y, y)})} \eta_{ks}(x) y^s.
\end{aligned}$$

One can check that $\frac{1}{2} \frac{\partial}{\partial y^j} \frac{\partial}{\partial y^k} G^i(x, y) = {}^L\Gamma^i{}_{jk}(x, y)$.

From the structure of these coefficients one can check (following a standard procedure, for instance in [34]) that there is a splitting of the tangent vector spaces $\mathbf{T}_u\mathbf{N}$ for each $u \in \mathbf{N}$. In addition, one can check that the connection coefficients are symmetric, ${}^L\Gamma^i{}_{jk} = {}^L\Gamma^i{}_{kj}$. The fact that the covariant derivative along the vertical directions is zero is an additional hypothesis used to make the covariant derivative unique. \square

Remark 1. In this interpretation of the Lorentz force equation, the role of the function F_A is not fundamental, since all the information is obtained directly from the differential equations (4.1.1).

Remark 2. We have imposed the restriction that the covariant derivatives along vertical directions are zero. However, there can be covariant derivatives which are compatible with the equations (4.1.1) but which have non-zero vertical covariant derivatives.

Proposition 4.4.3 *The following properties hold:*

1. *The Lorentz connection LD is invariant under gauge transformations $A \longrightarrow A + d\lambda$ of the locally smooth 1-form $A(x) = A_i(x)dx^i$.*
2. *Given a point $x \in \mathbf{M}$, LD admits a normal coordinate system centered at x which coincides with the normal coordinate system associated with ${}^\eta\nabla$ centered at x iff $\mathbf{F}(x) = 0$.*

Proof:

1. The first property is a consequence of the fact that all the geometric objects appearing in the connection coefficients ${}^L\Gamma^i{}_{jk}$ are gauge invariant.

2. If ${}^L D$ is affine, for any point $x \in \mathbf{M}$ there is a coordinate system where the connection coefficients are zero ${}^L \Gamma^i_{jk}(x, y) = 0$. This implies that ${}^L \Gamma^i_{jk} y^j y^k = 0$ at the point $x \in \mathbf{M}$. Due to the decomposition ${}^L \Gamma^i_{jk} = {}^\eta \Gamma^i_{jk} + T^i_{jk} + L^i_{jk}$ this is equivalent to

$$0 = ({}^\eta \Gamma^i_{jk} + T^i_{jk} + L^i_{jk}) y^j y^k, \quad \forall y \in \Sigma_x.$$

Since the transversality condition holds ($T^i_{jk} y^j y^k = 0$), one obtains

$$0 = ({}^\eta \Gamma^i_{jk} + L^i_{jk}) y^j y^k, \quad \forall y \in \Sigma_x.$$

Assume that there is a normal coordinate system centered at x for ${}^L \nabla$ and that this coordinate system coincides with the normal coordinate system associated with ${}^\eta \nabla$. Then at the point x , one has the relation

$$({}^L \Gamma^i_{jk} y^j y^k) = \mathbf{F}^i y^j y^k = 0, \quad \forall y \in \mathbf{T}_x \mathbf{M}.$$

This last condition is strong enough to imply $\mathbf{F} = 0$.

□

Remark. In the positive definite case, it is well known that the requirement that the Chern connection of a Randers space lives on the manifold \mathbf{M} is that the 1-form A must be parallel in the sense that ${}^\eta \nabla A = 0$, [24, chapter 11]. This is a stronger condition than $\mathbf{F} = dA = 0$. The parallel condition indicates that the structure (\mathbf{M}, F) is a generalization of the Berwald structure in Finsler geometry: a Berwald space is a Finsler space where the connection coefficients live on \mathbf{M} ; the closeness condition indicates that the space is Douglas [24, pg 304]; Douglas's spaces are such that they have the same geodesics as the underlying Riemannian metric η . One can easily move the proofs to the Lorentzian and indefinite category, if one adopt Asanov's framework.

Corollary 4.4.4 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space. Then the Lorentz force*

equation can be written as

$${}^L D_{\dot{\tilde{x}}} \dot{\tilde{x}} = 0,$$

where $x : \mathbf{I} \longrightarrow \mathbf{M}$ is a time-like curve parameterized with respect to the proper time of the Lorentzian metric η , \tilde{x} is the horizontal lift on \mathbf{N} and ${}^L \nabla$ is the non-linear connection determined by the system of differential equations (4.1.1).

Proof: A solution of the auto-parallel condition of the Lorentz connection defines a curve on \mathbf{N} given by $(x, y)(\tau) = (x(\tau), \dot{x}(\tau))$. Projecting this curve into \mathbf{M} by π , one obtains a curve $x(\tau)$ which is a solution of the Lorentz force equation. \square

4.4.3 The Lorentz connection on the pull-back bundle $\pi^* \mathbf{TM}$

We introduce the third framework, which will be directly used later to define the averaged connection. Given the non-linear connection ${}^L D$ on $\mathbf{TN} \longrightarrow \mathbf{N}$, there is a natural linear connection on the pull-back bundle $\pi^* \mathbf{TM} \longrightarrow \mathbf{N}$ that we denote by ${}^L \nabla$ characterized by the following:

Proposition 4.4.5 *The linear connection ${}^L \nabla$ on the pull-back bundle $\pi^* \mathbf{TM} \rightarrow \mathbf{N}$ is determined by the following structure equations,*

1. ${}^L \nabla$ on $\pi^* \mathbf{TM} \rightarrow \mathbf{N}$ is a symmetric connection,

$${}^L \nabla_{\tilde{X}} \pi^* Y - {}^L \nabla_{\tilde{Y}} \pi^* X - \pi^* [X, Y] = 0, \quad (4.4.2)$$

where $X, Y \in \Gamma \mathbf{TM}$, $\tilde{X}, \tilde{Y} \in \Gamma \mathbf{TN}$ are horizontal lifts of $X, Y \in \Gamma \mathbf{TM}$ to $\Gamma \mathbf{TN}$, with $\eta(X, X) > 0$ and $\eta(Y, Y) > 0$.

2. The covariant derivative along vertical directions of sections of $\pi^* \mathbf{TM}$ are zero,

$${}^L \nabla_{v_j} \pi^* e_k = 0, \quad (j, k = 0, \dots, n-1). \quad (4.4.3)$$

3. The covariant derivative along horizontal directions is given by the formula

$${}^L \nabla_{h_j} \pi^* e_k = {}^L \Gamma^i_{jk}(x, y) \pi^* e_i, \quad (i, j, k = 0, \dots, n-1). \quad (4.4.4)$$

4. By definition the covariant derivative of a function $f \in \mathcal{F}(\mathbf{N})$ is given by

$${}^L\nabla_{\hat{X}}f := \hat{X}(f), \quad \forall \hat{X} \in \mathbf{T}_u\mathbf{N}. \quad (4.4.5)$$

Proof: One can check by direct computation that the above relations define a *covariant derivative* on $\pi^*\mathbf{TM}$ and that they are self-consistent. A general covariant derivative can be expressed in terms of the connection 1-forms as

$$\omega^i{}_j := {}^L\Gamma^i{}_{jk}dx^k + {}^L\Upsilon^i{}_{jk}\delta y^k,$$

where we have used a local frame of 1-forms $\{dx^0, \dots, dx^{n-1}, \delta y^0, \dots, \delta y^{n-1}\}$. Since the covariant derivative of sections on $\pi^*\mathbf{TM}$ along vertical directions is zero, one obtains

$${}^L\Upsilon^i{}_{jk}\delta y^k = 0 \Rightarrow {}^L\Upsilon^i{}_{jk} = 0$$

at each point $(x, y) \in \mathbf{N}$. Since the torsion tensor is zero, one has

$${}^L\Gamma^i{}_{jk} = {}^L\Gamma^i{}_{kj}.$$

Therefore, we have to provide the rule for deriving sections along horizontal directions. Since the coefficients given by formula (4.4.1) are symmetric, this rule is consistent with the torsion-free condition. Finally, we require that the covariant derivative to be a local operator. This is satisfied by (4.4.5), which guaranties that ${}^L\nabla$ satisfies the Leibnitz rule. \square

Corollary 4.4.6 *Let \mathbf{M} , \mathbf{N} , $\pi^*\mathbf{TM}$ and ${}^L\nabla$ be as before. Then the auto-parallel curves of the linear Lorentz connection ${}^L\nabla$ are in one to one correspondence with the solutions of the Lorentz force equation,*

$${}^L\nabla_{\hat{x}}\pi^*\dot{x} = 0 \Leftrightarrow {}^LD_{\hat{x}}\dot{x} = 0, \quad \dot{x} = \frac{d\sigma}{d\tau}.$$

Proof: If in some coordinate system the Lorentz connection ${}^L\nabla$ has the connection coefficients ${}^L\Gamma^i{}_{jk}$, the auto-parallel condition is

$$\begin{aligned}
0 &= \left(\pi^* \left({}^L\nabla_{\frac{dx(\tau)}{d\tau}} \pi^* \frac{dx(\tau)}{d\tau} \right) \right)^i = \left(\frac{d^2 x^i(\tau)}{d\tau^2} + {}^L\Gamma^i{}_{jk}(x, \frac{dx(\tau)}{d\tau}) \frac{dx^j(\tau)}{d\tau} \frac{dx^k(\tau)}{d\tau} \right) = \\
&= \left(\frac{d^2 x^i(\tau)}{d\tau^2} + \left(\eta \Gamma^i{}_{kj} - \eta^{il}(dA)_{lm} \frac{1}{2(\eta(\frac{dx(\tau)}{d\tau}, \frac{dx(\tau)}{d\tau}))^{3/2}} \eta_{js} \frac{dx^s(\tau)}{d\tau} \frac{dx^m(\tau)}{d\tau} \eta_{kp} \frac{dx^p(\tau)}{d\tau} + \right. \right. \\
&+ \eta^{il}(dA)_{lm} \frac{1}{2(\sqrt{\eta(\frac{dx(\tau)}{d\tau}, \frac{dx(\tau)}{d\tau}))} \eta_{jk} \frac{dx^m(\tau)}{d\tau} + \eta^{il}(dA)_{lk} \frac{1}{2(\sqrt{\eta(\frac{dx(\tau)}{d\tau}, \frac{dx(\tau)}{d\tau}))} \eta_{js} \frac{dx^s(\tau)}{d\tau} + \\
&\left. \left. + \eta^{il}(dA)_{lj} \frac{1}{2(\sqrt{\eta(\frac{dx(\tau)}{d\tau}, \frac{dx(\tau)}{d\tau}))} \eta_{ks} \frac{dx^s(\tau)}{d\tau} \right) \frac{dx^j(\tau)}{d\tau} \frac{dx^k(\tau)}{d\tau} \right).
\end{aligned}$$

Using $\eta(\frac{dx(\tau)}{d\tau}, \frac{dx(\tau)}{d\tau}) = 1$, the above expression simplifies to

$$\begin{aligned}
0 &= \left(\pi^* \left({}^L\nabla_{\frac{dx(\tau)}{d\tau}} \frac{dx(\tau)}{d\tau} \right) \right)^i = \\
&= \left(\frac{d^2 x^i(\tau)}{d\tau^2} + \left(\eta \Gamma^i{}_{kj} + \eta^{il}(dA)_{lk} \frac{1}{2} \eta_{js} \frac{dx^s(\tau)}{d\tau} + \eta^{il}(dA)_{lj} \frac{1}{2} \eta_{ks} \frac{dx^s(\tau)}{d\tau} \right) \frac{dx^j(\tau)}{d\tau} \frac{dx^k(\tau)}{d\tau} \right).
\end{aligned}$$

This is the Lorentz force equation (4.1.1). \square

We need to translate from the non-linear connection ${}^L\nabla$ on \mathbf{TN} to the linear connection on $\pi^*\mathbf{TM}$ because we will consider the averaged connection, which was defined in [22] and in *chapter 3*.

4.5 The averaged Lorentz connection

Given the linear connection ${}^L\nabla$ on the bundle $\pi^*\mathbf{TM} \rightarrow \Sigma$ we can obtain an associated averaged connection using the theory described in *section 3.3*.

Usually the measure is given as $f(x, y) dvol(x, y)$. The function $f(x, y)$ must be gauge invariant and such that the low order moments are finite. The volume form is

$$dvol(x, y) = \sqrt{-\det \eta} \frac{1}{y^0} dy^1 \wedge \cdots \wedge dy^{n-1}, \quad y^0 = y^0(x^0, \dots, x^{n-1}, y^1, \dots, y^{n-1}).$$

Although the dimension of the manifold Σ is $2n - 1$, we will use the extrinsic formalism explained in *chapter 2*. This in particular means that all Latin indices run from 0 to $n - 1$, if nothing else is stated. Then one can prove the following

Proposition 4.5.1 *The averaged connection of the Lorentz connection ${}^L\nabla$ on the pull-back bundle $\pi^*\mathbf{TM} \rightarrow \Sigma$ is an affine, symmetric connection on \mathbf{M} . The connection coefficients are given by the formula*

$$\begin{aligned} \langle {}^L\Gamma^i_{jk} \rangle = & \eta\Gamma^i_{jk} + (\mathbf{F}^i{}_j \langle \frac{1}{2}y^m \rangle \eta_{mk} + \mathbf{F}^i{}_k \langle \frac{1}{2}y^m \rangle \eta_{mj}) + \\ & + \mathbf{F}^i{}_m \frac{1}{2} (\langle y^m \rangle \eta_{jk} - \eta_{js}\eta_{kl} \langle y^m y^s y^l \rangle). \end{aligned} \quad (4.5.1)$$

Each of the integrations is equal to the y -integration along the fiber,

$$\begin{aligned} \text{vol}(\Sigma_x) &= \int_{\Sigma_x} f(x, y) d\text{vol}(x, y), \quad \langle y^i \rangle := \frac{1}{\text{vol}(\Sigma_x)} \int_{\Sigma_x} y^i f(x, y) d\text{vol}(x, y), \\ \langle y^m y^s y^l \rangle &:= \frac{1}{\text{vol}(\Sigma_x)} \int_{\Sigma_x} y^m y^s y^l f(x, y) d\text{vol}(x, y). \end{aligned}$$

Proof: Equation (4.5.1) follows easily from the definition of the averaged connection by linearity. We only need to prove that $\langle y^i \rangle$ and the other moments are given by the corresponding integrals and the identity operator $Id : \pi^*\mathbf{TM} \rightarrow \pi^*\mathbf{TM}$, $(x, y) \mapsto (x, y)$:

$$\langle y^i \rangle = \frac{1}{\text{vol}(\Sigma_x)} \int_{\Sigma_x} \pi_2 Id y^i \pi^*(y^i) f(x, y) d\text{vol}(x, y) = \frac{1}{\text{vol}(\Sigma_x)} \int_{\Sigma_x} y^i f(x, y) d\text{vol}(x, y)$$

and similarly for other moments. Note that since $y \in \Sigma_x$ $\eta(y, y) = 1$, the factors $\sqrt{\eta(y, y)}$ do not appear in the connection coefficients. \square

Remarks.

1. If we consider the bundle $\pi^*\mathbf{TM} \rightarrow \mathbf{N}$, the coefficients of the averaged Lorentz connection are

$$\begin{aligned}
\langle {}^L\Gamma \rangle^i{}_{kj} &= \eta \Gamma^i{}_{kj}(x) - \eta^{il}(x)(dA)_{lm}(x) \langle \frac{1}{2(\eta(y,y))^{3/2}} \eta_{js}(x) y^s y^m \eta_{kp}(x) y^p \rangle + \\
&+ \eta^{il}(x)(dA)_{lm}(x) \langle \frac{1}{2(\sqrt{\eta(y,y)})} \eta_{jk}(x) y^m \rangle + \eta^{il}(x) \langle (dA)_{lk}(x) \frac{1}{2(\sqrt{\eta(y,y)})} \eta_{js} y^s \rangle + \\
&+ \eta^{il}(x)(dA)_{lj}(x) \langle \frac{1}{2(\sqrt{\eta(y,y)})} \eta_{ks}(x) y^s \rangle .
\end{aligned}$$

In this case, the averaged connection of higher moments of the distribution function.

2. Formula (4.5.1) holds in any local natural coordinate system.

The following proposition enumerates basic properties of the averaged Lorentz connection. The proof is straightforward; one needs to check the properties of the coefficients given by the formula (4.5.1):

Proposition 4.5.2 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space, $f : \Sigma \longrightarrow \mathbf{R}$ non-negative with compact support in each Σ_x and $\langle {}^L\nabla \rangle$ be the averaged Lorentz connection. Then*

1. *$\langle {}^L\nabla \rangle$ is an affine, symmetric connection on \mathbf{M} . Therefore, for any point $x \in \mathbf{M}$, there is a normal coordinate system such that $\langle {}^L\Gamma^i{}_{jk} \rangle(x) = 0$.*
2. *$\langle {}^L\nabla \rangle$ is determined by the first, second and third moments of the distribution function $f(x, y)$.*

Remark. While the first property is a general property of the averaged connection, the second one is a specific property of the averaged Lorentz connection, which results because we are considering trajectories whose velocity vectors y in the unit hyperboloid Σ . Also note that $\langle {}^L\nabla \rangle$ does not preserve the norm $\eta_{ij}(x)y^i y^j$ for arbitrary $y \in \mathbf{T}_x \mathbf{M}$: $\langle {}^L\nabla \rangle_Z (\eta_{ij}(x)y^i y^j) \neq 0$ for an arbitrary $Z \in \mathbf{T}_x \mathbf{M}$.

4.6 Comparison between the geodesics of ${}^L\nabla$ and

$$< {}^L\nabla >$$

4.6.1 Basic geometry in the space of connections

Let us consider the Lorentzian manifold (\mathbf{M}, η) with signature $(+, -, \dots, -)$ and \mathbf{M} n -dimensional. Let us choose on \mathbf{M} a time-like vector field U normalized such that $\eta(U, U) = 1$. Then one can define the Riemannian metric $\bar{\eta}_U$ [36]

$$\bar{\eta}_U(X, Y) := -\eta(X, Y) + 2\eta(X, U)\eta(Y, U). \quad (4.6.1)$$

$\bar{\eta}_U$ determines a Riemannian metric on the vector space $\mathbf{T}_x\mathbf{M}$ that we also denote by $\bar{\eta}_U$ and that in local coordinates can be expressed as $\bar{\eta}_U = \bar{\eta}_{ij}(x) dy^i \otimes dy^j$. Note that the metric $\bar{\eta}_U$ on \mathbf{M} is in local coordinates $\bar{\eta}_U = \bar{\eta}_{ij}(x) dx^i \otimes dx^j$. We hope that the meaning of the symbols of the type $\bar{\eta}_U$ is clear from the context.

The pair $(\mathbf{T}_x\mathbf{M}, \bar{\eta}_U)$ is a Riemannian manifold. The Riemannian metric $\bar{\eta}_U$ induces a distance function $d_{\bar{\eta}_U}$ on the manifold $\mathbf{T}_x\mathbf{M}$,

$$d_{\bar{\eta}_U} : \mathbf{T}_x\mathbf{M} \times \mathbf{T}_x\mathbf{M} \longrightarrow \mathbf{R}$$

$$(X, Y) \mapsto \inf \left\{ \int_0^1 \sqrt{\bar{\eta}_U(\dot{\hat{\sigma}}, \dot{\hat{\sigma}})} d\tau, \hat{\sigma} : \mathbf{I} \longrightarrow \mathbf{T}_x\mathbf{M}, \hat{\sigma}(0) = X, \hat{\sigma}(1) = Y \right\}.$$

We will say that the time-like vector field U defines a field of observer [60, pg 45].

We assume that $f(x, y)$ has compact support on each unit hyperboloid Σ_x . Then the diameter of the distribution $f_x = f(x, \cdot) : \Sigma_x \longrightarrow \mathbf{R}$; $(x, y) \mapsto f(x, y)$ is $\alpha_x := \sup\{d_{\bar{\eta}}(y_1, y_2) \mid y_1, y_2 \in \text{supp}(f_x)\}$. We define the parameter $\alpha := \sup\{\alpha_x, x \in \mathbf{K}\}$, with $\mathbf{K} \subset \mathbf{M}$ a compact domain of the space-time \mathbf{M} . We will restrict always our considerations to \mathbf{K} . Note that α (and other parameters) can depend on \mathbf{K} .

Let us fix the coordinate system (x, y) on \mathbf{N} . Let us denote the space of linear connections on \mathbf{TN} by $\nabla_{\mathbf{N}}$. This space is a finite dimensional manifold whose points are coordinated by the set of functions $\{\Gamma^i_{jk}(x, y), i, j, k = 0, 1, \dots, n-1\}$. We will introduce a distance function on $\nabla_{\mathbf{N}}$. First, we recall the definition of the norm of

an operator.

Definition 4.6.1 *Given a linear operator $A_x : \mathbf{T}_x \mathbf{M} \longrightarrow \mathbf{T}_x \mathbf{M}$ in the finite dimensional normed linear space $(\mathbf{T}_x \mathbf{M}, \|\cdot\|_{\bar{\eta}_Z})$, its operator norm is defined by*

$$\|A\|_{\bar{\eta}_Z}(x) := \sup \left\{ \frac{\|A(y)\|_{\bar{\eta}_Z}(x)}{\|y\|_{\bar{\eta}_Z}}, y \in \mathbf{T}_x \mathbf{M} \setminus \{0\} \right\}.$$

The norm $\|\cdot\|_{\bar{\eta}_Z}$ is constructed from the Lorentzian metric η using a particular local time-like vector Z through the definition (4.6.1). We will later specify a vector field Z which will be of special interest for our purposes.

Proposition 4.6.2 *On the space $\nabla_{\mathbf{N}}$, there is a distance function. The distance between two points ${}^1\nabla, {}^2\nabla \in \nabla_{\mathbf{N}}$ is given by*

$$d_{\bar{\eta}_Z}({}^1\nabla, {}^2\nabla)(x) := \sup \left\{ \frac{\sqrt{\bar{\eta}_Z(x)({}^1\nabla_X X - {}^2\nabla_X X, {}^1\nabla_X X - {}^2\nabla_X X)}}{\sqrt{\bar{\eta}_Z(X, X)}}, \right. \\ \left. X \in \Gamma \mathbf{T} \mathbf{N}, \quad {}^1\nabla, {}^2\nabla \in \nabla_{\mathbf{N}} \right\} \quad (4.6.2)$$

for a Riemannian metric (4.6.1) constructed using the local time-like vector field Z .

Proof: The function (4.6.2) is symmetric and non-negative. The distance between two arbitrary connections is zero iff

$$\sqrt{\bar{\eta}_Z({}^1\nabla_X X - {}^2\nabla_X X, {}^1\nabla_X X - {}^2\nabla_X X)} = 0$$

for all $X \in \Gamma \mathbf{T} \Sigma$. This happens iff ${}^1\nabla_X X = {}^2\nabla_X X$ for any $X \in \Gamma \mathbf{T} \Sigma$. The triangle inequality also holds, from the triangle inequality for $\bar{\eta}$. This reads

$$d_{\bar{\eta}_Z}(\nabla_1, \nabla_3) = \sup \left\{ \frac{\sqrt{\bar{\eta}_Z({}^1\nabla_X X - {}^3\nabla_X X, {}^1\nabla_X X - {}^3\nabla_X X)}}{\sqrt{\bar{\eta}_Z(X, X)}} \right\} \leq \\ \leq \sup \left\{ \frac{\sqrt{\bar{\eta}_Z({}^1\nabla_X X - {}^2\nabla_X X, {}^1\nabla_X X - {}^2\nabla_X X)}}{\sqrt{\bar{\eta}_Z(X, X)}} \right\} + \\ + \sup \left\{ \frac{\sqrt{\bar{\eta}_Z({}^2\nabla_X X - {}^3\nabla_X X, {}^2\nabla_X X - {}^3\nabla_X X)}}{\sqrt{\bar{\eta}_Z(X, X)}} \right\} \leq$$

$$\leq d_{\bar{\eta}_Z}({}^1\nabla, {}^2\nabla) + d_{\bar{\eta}_Z}({}^2\nabla, {}^3\nabla).$$

□

Recall that for an arbitrary 1-form ω we denote by $\omega^\sharp := \eta^{-1}(\omega, \cdot)$ the vector obtained by duality, using the Lorentzian metric η . Similarly, given a vector field X over \mathbf{M} , one can define the dual one form $X^\flat := \eta(X, \cdot)$; $\iota_X \omega$ is the inner product of the vector X with the form ω . Also recall that the difference between two connections is a tensor. In this sense, one can consider the difference between connections on the pull-back bundle $\pi^*\mathbf{TM}$ given by the Lorentz connection ${}^L\nabla$ and the pull-back connection of the averaged connection $\pi^* \langle {}^L\nabla \rangle$. Then it makes sense to take the difference ${}^L\nabla - \pi^* \langle {}^L\nabla \rangle$, which is a tensor along π . We use the *hat*-notation to indicate integrated variables. For instance, $\langle \hat{y} \rangle$ means an integration operation, where the variable integrated is \hat{y} . Let us consider a local frame $\{e_0, \dots, e_{n-1}\}$ on \mathbf{TM} .

Proposition 4.6.3 *Let $f(x, y)$ be the one-particle probability distribution function such that each function f_x has compact and connected support $\text{supp}(f_x) \subset \Sigma_x$. Then*

$$({}^L\nabla_y y - \pi^* \langle {}^L\nabla \rangle_y y)(x) = -(\iota_\delta \mathbf{F})^\sharp(x) \cdot (\iota_y(\delta))(x, y) + \mathcal{O}_2(\delta^2(y))(x, y) + \mathcal{O}_3(\delta^3(y))(x, y), \quad (4.6.3)$$

where $\delta(x, y) = \langle \hat{y} \rangle(x) - y$ does not depend on the 2-form \mathbf{F} . The tensors $\mathcal{O}_i(x, y)$ are given by the following expressions:

$$\begin{aligned} \mathcal{O}_2(\delta^2(y))(x, y) = \frac{1}{2} \mathbf{F}^i{}_m \Big(\langle \hat{y}^m \rangle(x) \delta^s(x, y) \delta^l(x, y) + \langle \hat{y}^m \rangle \langle \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \\ + 2 \langle \hat{y}^l \rangle(x) \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \rangle \Big) \eta_{sj} \eta_{lk} y^j y^k \pi^* e_i, \end{aligned} \quad (4.6.4)$$

$$\mathcal{O}_3(\delta^3(y))(x, y) = \frac{1}{2} \mathbf{F}^i{}_m \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \eta_{sj} \eta_{lk} y^j y^k \pi^* e_i. \quad (4.6.5)$$

Proof: From the expressions for the connection coefficients,

$$\pi^* \langle {}^L\nabla \rangle_y y - {}^L\nabla_y y = \frac{1}{2} \left(\mathbf{F}^i{}_j(x) (\langle \hat{y}^m \rangle(x) - y^m) \eta_{mk} + \mathbf{F}^i{}_k(x) (\langle \hat{y}^m \rangle(x) - y^m) \eta_{mj} \right) +$$

$$+\mathbf{F}^i_m(x)(\langle \hat{y}^m \rangle(x) - y^m)\eta_{jk} - \eta_{js}\eta_{kl}(\langle \hat{y}^m \hat{y}^s \hat{y}^l \rangle(x) - y^m y^s y^l))y^j y^k \pi^* e_i,$$

since $\eta_{ij}y^i y^j = 1$. Here $\{e_i, i = 0, \dots, n-1\}$ is an arbitrary frame unless otherwise is indicated. The difference between the two connections can be expressed in terms of the following tensors:

$$\delta^m(x, y) = \langle \hat{y}^m \rangle(x) - y^m, \quad \delta^{msl}(x, y)y_s y_l = \langle \hat{y}^m \hat{y}^s \hat{y}^l \rangle(x) \eta_{sj} \eta_{lk} y^j y^k - y^m$$

and is given by the following expression:

$$\begin{aligned} \left(\pi^* \langle {}^L \nabla \rangle_y y - {}^L \nabla_y y \right)(x) &= \frac{1}{2} \left(\mathbf{F}^i_j(x) (\delta^m(x, y)) \eta_{mk} + \mathbf{F}^i_k(x) (\delta^m(x, y)) \eta_{mj} + \right. \\ &\quad \left. + \mathbf{F}^i_m(x) (\delta^m(x, y) \eta_{jk} - \eta_{js} \eta_{kl} \delta^{msj}(x, y)) \right) y^j y^k \pi^* e_i = \\ &= \left(\mathbf{F}^i_j y^j \delta^k(x, y) y_k + \mathbf{F}^i_m(x) (\delta^m(x, y) - y_j y_k \delta^{mkj}(x, y)) \right) \pi^* e_i, \end{aligned}$$

where $y_j = \eta_{jk} y^k$.

In the above subtractions the second contribution is of the same order in $\delta(y)$ as the first one. To show this, recall from the definitions

$$\delta^{msl}(x, y)y_s y_l = \langle \hat{y}^m(x) \hat{y}^s(x) \hat{y}^l(x) \rangle \eta_{sj} \eta_{lk} y^j y^k - y^m.$$

Then we can use the following relations:

$$\hat{y}^s = \langle \hat{y}^s \rangle(x) - \delta^s(x, \hat{y}).$$

One substitutes this relation and taking into account that $\langle \delta^k(x, \hat{y}) \rangle = 0$, one gets

$$\begin{aligned} \delta^{msl}(x, y)y_s y_l &= \left(\langle \hat{y}^m \rangle(x) \langle \hat{y}^s \rangle(x) \langle \hat{y}^l \rangle(x) + \langle \hat{y}^m \rangle(x) \langle \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle + \right. \\ &\quad \left. + \langle \hat{y}^l \rangle(x) \langle \delta^s(x, \hat{y}) \delta^m(x, \hat{y}) \rangle + \langle \hat{y}^s \rangle(x) \langle \delta^m(x, \hat{y}) \delta^l(x, \hat{y}) \rangle - \right. \\ &\quad \left. - \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \right) \eta_{sj} \eta_{lk} y^j y^k - y^m. \end{aligned}$$

Now we use a similar relation to go further in the calculation. Let us write

$$y^s = \langle \hat{y}^s \rangle (x) - \delta^s(x, y).$$

We introduce these expressions in the calculation of $\delta^{msl} y_s y_l$:

$$\begin{aligned} \delta^{msl}(x, y) y_s y_l = & \left(\langle \hat{y}^m \rangle (x) (y^l + \delta^l(x, y)) (y^s + \delta^s(x, y)) + \langle \hat{y}^m \rangle (x) \langle \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle + \right. \\ & + (y^l + \delta^l(x, y)) \langle \delta^s(x, \hat{y}) \delta^m(x, \hat{y}) \rangle + (y^s + \delta^s(x, y)) \langle \delta^m(x, \hat{y}) \delta^l(x, \hat{y}) \rangle - \\ & \left. - \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \right) \eta_{sj} \eta_{lk} y^j y^k - y^m. \end{aligned}$$

Using again the fact that $y^l \eta_{lk} y^k = 1$ we get (again using $\langle \hat{y}^m \rangle = y^m + \delta^m(x, y)$ to recombine the first and last term),

$$\delta^{msl}(x, y) y_s y_l = \delta^m(x, y) + 2 \langle \hat{y}^m \rangle (x) \delta^s(y) \eta_{sj} y^j + 2\mathcal{O}_2^m(\delta^2) + 2\mathcal{O}_3^m(\delta^3),$$

the tensors \mathcal{O}_2 and \mathcal{O}_3 are given by the formulae

$$\begin{aligned} \mathcal{O}_2^i(\delta^2(y))(x, y) &= \frac{1}{2} \mathbf{F}^i_m \left(\langle \hat{y}^m \rangle (x) \delta^s(x, y) \delta^l(x, y) + \langle \hat{y}^m \rangle \langle \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \right. \\ &\quad \left. + 2 \langle \hat{y}^l \rangle (x) \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \rangle \right) \eta_{sj} \eta_{lk} y^j y^k, \\ \mathcal{O}_3^i(\delta^3(y))(x, y) &= \frac{1}{2} \mathbf{F}^i_m \langle \delta^m(x, \hat{y}) \delta^s(x, \hat{y}) \delta^l(x, \hat{y}) \rangle \eta_{sj} \eta_{lk} y^j y^k. \end{aligned}$$

The *transversal contribution* to the difference between the connections is given by:

$$\begin{aligned} \frac{1}{2} \mathbf{F}^i_m(x) (\delta^m(x, y) \eta_{jk} - \eta_{js} \eta_{kl} \delta^{msj}(x, y) y^j y^k) &= \frac{1}{2} \mathbf{F}^i_m(x) (\delta^m(x, y) - \delta^m(x, y) \\ &\quad - 2 \langle \hat{y}^m \rangle (x) \delta^s(x, y) \eta_{sj} y^j) - (\mathcal{O}_2^m(\delta^2) - \mathcal{O}_3^m(\delta^3)) \mathbf{F}^i_m = \\ &= -\mathbf{F}^i_m(x) \left(\langle \hat{y}^m \rangle (x) \delta^s(x, y) \eta_{sj} y^j - \frac{1}{2} \mathcal{O}_2^m(\delta^2) - \frac{1}{2} \mathcal{O}_3^m(\delta^3) \right) = \\ &= -\mathbf{F}^i_m(x) \left(\langle \hat{y}^m \rangle (x) \delta^s(x, y) y_s - \frac{1}{2} \mathcal{O}_2^m(\delta^2) - \frac{1}{2} \mathcal{O}_3^m(\delta^3) \right) = \end{aligned}$$

The *longitudinal contribution* is

$$\frac{1}{2}(\mathbf{F}^i{}_j(x)(\langle \hat{y}^m \rangle(x) - y^m)\eta_{mk} + \mathbf{F}^i{}_k(x)(\langle \hat{y}^m \rangle(x) - y^m)\eta_{mj}) y^j y^k = \mathbf{F}^i{}_j(x) y^j \delta^k(x, y) y_k.$$

Adding together the *longitudinal* and *transversal* contributions and taking into account the formula $\delta^m(x, y) = \langle \hat{y}^m \rangle(x) - y^m$ we get the following expression:

$$\begin{aligned} & (\mathbf{F}^i{}_j y^j)(\delta^k(x, y) y_k) - \left(\mathbf{F}^i{}_m(x) (\langle \hat{y}^m \rangle(x) \delta^s(x, y) \eta_{sj} y^j + \frac{1}{2} \mathcal{O}_2(\delta^2) + \frac{1}{2} \mathcal{O}_3(\delta^3)) \right) = \\ & = -(\mathbf{F}^i{}_m \delta^m(x, y)) (\delta^k(x, y) y_k) - \left(\frac{1}{2} \mathcal{O}_2^m(\delta^2) + \frac{1}{2} \mathcal{O}_3^m(\delta^3) \right) \mathbf{F}^i{}_m. \end{aligned}$$

□

Let us consider a frame $\{e_i, i = 0, \dots, n-1\}$ such that $\bar{\eta}_Z$ is diagonal at the point $x \in \mathbf{M}$ in this frame. After calculating the distance of the connections using the formula (4.6.3), the leading term in δ is quadratic. Also recall that in *section 1.3* we have fixed our units of energy and momentum in such a way that they are given by dimensionless numbers.

Remark. Given a norm $\bar{\eta}_Z$ on $\mathbf{T}_x \mathbf{M}$, one can define an induced distance on $\pi_1^{-1}(u)$, $u \in \pi^{-1}(x)$. This norm is just defined as $d_{\bar{\eta}_Z}(\xi, \zeta) := d_{\bar{\eta}_Z}(\pi_1(\xi), \pi_1(\zeta))$, $\zeta, \xi \in \pi_1^{-1}(u)$.

Proposition 4.6.4 *Let $(\mathbf{M}, \eta, [A])$ and ${}^L\nabla$ be as before and assume that f_x has compact and connected support for each fixed $x \in \mathbf{M}$ and that $\alpha := \sup\{\alpha_x, x \in \mathbf{M}\} < 1$. Then the following holds,*

$$d_{\bar{\eta}_Z}(\pi^* \langle {}^L\nabla \rangle, {}^L\nabla)(x) \leq \|\mathbf{F}\|_{\bar{\eta}_Z}(x) C(x) \alpha^2 + 2C_2^2(x) \alpha^2 (1 + \alpha) + C_3^3(x) \alpha^3 (1 + \alpha), \quad (4.6.6)$$

with $C(x), C_2(x), C_3(x)$ being functions depending only on x with value of the order of unity.

Proof: Let $\{\pi^* e_0, \dots, \pi^* e_{n-1}\}$ be a local orthonormal frame for the induced fiber metric on the pull-back bundle from the Riemannian metric $\bar{\eta}$. From equation

(4.6.3) one obtains

$$\begin{aligned} \| {}^L \nabla_y y - \pi^* < \hat{y} >_y y \|_{\bar{\eta}_Z} &= \| \mathbf{F}^i_j(x) \delta^j(x, y) \delta^k(x, y) y_k \pi^* e_i + \mathcal{O}_2(\delta^2)(x, y) + \mathcal{O}_3(\delta^3)(x, y) \|_{\bar{\eta}_Z} \leq \\ &\leq \| \mathbf{F}^i_j \delta^j(x, y) \delta^k(x, y) y_k e_i \|_{\bar{\eta}_Z} + \| \mathcal{O}_2(\delta^2)(x, y) \|_{\bar{\eta}_Z} + \| \mathcal{O}_3(\delta^3)(x, y) \|_{\bar{\eta}_Z}. \end{aligned}$$

Each of these three terms can be bounded.

Recall that we are using a local frame such that $\|e_i\|_{\bar{\eta}_Z} = 1$. Then we can bound the first term in the following way:

$$\| \mathbf{F}^i_j \delta^j(x, y) \delta^k(x, y) y_k e_i \|_{\bar{\eta}_Z} \leq \| \mathbf{F} \|_{\bar{\eta}_Z} \cdot \| \delta(x, y) \|_{\bar{\eta}_Z} \cdot | \delta^k(x, y) y_k |.$$

For a fixed x the support of the distribution function $f(x, y)$ is compact and connected, thus one can write the decomposition $< \hat{y} > (x) = \epsilon(x) + z(x)$ with the property $z(x) \in \text{supp}(f_x)$. In the case that η is the Minkowski metric one can check that $\| \epsilon(x) \|_{\bar{\eta}_Z} \leq \alpha(x) \leq \alpha$ by geometric inspection. This bound of $\epsilon(x)$ follows from the shape of the unit hyperboloid and is proved in the following way. First, note that the domain $\widehat{\Sigma}_x := \{y \in \mathbf{T}_x \mathbf{M} \mid \eta(y, y) \geq 1, y^0 > 0\}$ is a convex set with respect to $\bar{\eta}$. Indeed, we note that $\partial(\widehat{\Sigma}_x) = \Sigma_x$ and that $< \hat{y} > (x) \in \widehat{\Sigma}_x$. Secondly, each $(\widehat{\Sigma}_x, \bar{\eta}_{Z_x})$ is a Riemannian manifold, with $\bar{\eta}_{Z_x} := (\bar{\eta}_Z)_{ij}(x) dy^i \otimes dy^j$. Therefore, we can use the standard definition of center of mass [48], in this case with a measure given by $f(x, y, s) d\text{vol}(x, y)$. The function $f(x, y, s)$ is such that

$$\int_0^1 ds f(x, y, s) = f(x, y),$$

s is the parameter of the line connecting $y \in \text{supp}(f_x)$ with $< \hat{y} > (x)$. Let us denote by $\widehat{\text{supp}}(f_x)$ the convex hull of $\text{supp}(f_x)$. By construction $< \hat{y} > (x) \in \widehat{\text{supp}}(f_x)$. One can check that $< \hat{y} > (x)$ is the center of mass of the convex set $\widehat{\text{supp}}(f_x)$ [48].

We have the following bound

$$\| \delta(x, y) \|_{\bar{\eta}_Z} \leq \| < \hat{y} > (x) - y \|_{\bar{\eta}_Z} \leq \| \epsilon(x) + z(x) - y \|_{\bar{\eta}_Z} \leq$$

$$\leq \|\epsilon(x)\|_{\bar{\eta}_Z} + \|z(x) - y\|_{\bar{\eta}_Z} \leq \alpha + \alpha = 2\alpha.$$

For the third factor, one has the following bound ($\delta(x, y) = \langle y \rangle - y$)

$$\begin{aligned} |\delta^k(x, y)y_k| &= |\langle \hat{y}^k \rangle(x)y_k - 1| = |\langle \hat{y}^k \rangle(x)(y_k - \langle \hat{y}^k \rangle(x) + \langle \hat{y}^k \rangle(x)) - 1| \leq \\ &\leq |\langle \hat{y}^k \rangle(x)(y_k - \langle \hat{y}^k \rangle(x))| + |\langle \hat{y}^k \rangle(x) - 1|. \end{aligned}$$

Using Cauchy-Schwarz inequality for $\bar{\eta}_Z$ we obtain

$$\begin{aligned} |\delta^k(x, y)y_k| &\leq \|\langle \hat{y}^k \rangle(x)\|_{\bar{\eta}_Z} \|y_k - \langle \hat{y}^k \rangle(x)\|_{\bar{\eta}_Z} + |\langle \hat{y}^k \rangle(x) - 1| \leq \\ &\leq \|\langle \hat{y}^k \rangle(x)\|_{\bar{\eta}_Z} \alpha + |\langle \hat{y}^k \rangle(x) - 1| \leq \sqrt{1 + \|\epsilon\|_{\bar{\eta}_Z}^2} \alpha + (\sqrt{1 + \|\epsilon(x)\|_{\bar{\eta}_Z}^2} - 1) \leq \\ &\leq \sqrt{1 + \alpha} \alpha + (\sqrt{1 + \alpha} - 1) \leq (1 + \alpha)\alpha + (1 + \alpha - 1) = 2\alpha + \alpha^2. \end{aligned}$$

Therefore, one obtains

$$\|(\mathbf{F}^i_j \delta^j(x, y))(\delta^k(x, y)y_k) e_i\|_{\bar{\eta}_Z} \leq \|\mathbf{F}\|_{\bar{\eta}_Z}(x) C(x) \alpha^2 + \mathcal{O}(\alpha^4).$$

The function $C(x)$ in equation (4.6.6) is bounded by the constant 4 in the coordinate frame determined by the vector field Z . This bound is universal, independent of the Lorentzian metric η , the vector field Z and it has a geometric origin.

Using homogeneity properties on the variable y , one can see that the following relations hold:

$$\|\mathcal{O}_2(\delta^2)\|_{\bar{\eta}_Z} \leq C_2^2(x) \alpha^2 (1 + B_2(x) \alpha), \quad (4.6.7)$$

and

$$\|\mathcal{O}_3(\delta^3)\|_{\bar{\eta}_Z} \leq C_3^3(x) \alpha^3 (1 + B_3(x) \alpha). \quad (4.6.8)$$

The functions $C_i(x)$ depend on the particular shape of the support of the distribution function f and on the curvature of the metric η . Using geometric arguments (and in particular compactness and connectedness of the $\text{supp}(f(x))$) in \mathbf{K} , one can bound these functions in terms of α in a similar way as we did for $C(x)$. The constants are of order 1 because this was the case for $C(x)$ and there are no new *divergence*

factors in the functions $B_i(x)$ and $C_i(x)$. \square

Corollary 4.6.5 *Let $(\mathbf{M}, \eta, [A])$ be a (semi)-Randers space. Let us consider a compact domain $\mathbf{K} \subset \bigsqcup_{x \in \mathbf{M}} \text{supp}(f_x)$ compact, with ${}^L\nabla$ and $\langle {}^L\nabla \rangle$ as before. Then there is a global bound:*

$$d_{\bar{\eta}_Z}({}^L\nabla, \pi^* \langle {}^L\nabla \rangle)(x) \leq C \|\mathbf{F}\|_{\bar{\eta}_Z} \alpha^2 + 2C_2^2 \alpha^2 (1 + B_2 \alpha) + C_3^3 \alpha^3 (1 + B_3 \alpha), \quad \forall x \in \mathbf{M}.$$

where the constants C, C_2, C_3, B_2, B_3 are bounded by a constant of order 1.

Proof: It follows from *proposition* (4.6.4) and compactness of the domain $\mathbf{K} \subset \bigsqcup_{x \in \mathbf{M}} \text{supp}(f_x)$ that we are considering. \square

4.6.2 Comparison between the geodesics of ${}^L\nabla$ and $\pi^* \langle {}^L\nabla \rangle$

There are several ways of defining the energy of a bunch of particles. We have chosen one which will be useful for our comparison results. We define the energy function E of a distribution f to be the real function

$$E : \mathbf{M} \longrightarrow \mathbf{R}$$

$$x \mapsto E(x) := \inf \{y^0, y \in \text{supp}(f_x)\}, \quad (4.6.9)$$

where y^0 is the 0-component of a tangent vector of a possible trajectory of a charged point particle, measured in the laboratory coordinate frame. The name *energy* for this function is deserved because of the choice of the units that we have adopted, even if energy is a function on the co-tangent bundle $\mathbf{T}^*\mathbf{M}$ instead of the bundle \mathbf{TM} , where the velocities are defined.

Let us restrict our attention to the case that the Lorentzian metric is the Minkowski metric in dimension n . We can define $\theta^2(t) = \vec{y}^2(t) - \langle \vec{y} \rangle^2(t)$ and $\bar{\theta}^2(t) = \langle \vec{y} \rangle^2(t) - \vec{\bar{y}}^2(t)$. Here $\vec{y}(t)$ is the spatial component of the velocity tangent vector field along a solution of the Lorentz force equation and $\vec{\bar{y}}(t)$ is the spatial component of the tangent vector field along a solution of the averaged Lorentz force equation,

with both solutions having the same initial conditions. The spatial components are defined respect the observer $Z = \frac{\partial}{\partial t}$. We will call this observer the laboratory frame. Since we are fixing this frame, the corresponding Riemannian metric $\|\cdot\|_{\bar{\eta}_Z}$ is denoted simply by $\|\cdot\|_{\bar{\eta}}$, simplifying the notation. The maximal values of these quantities on the compact domain \mathbf{K} of the space-time manifold \mathbf{M} are denoted by θ^2 and $\bar{\theta}^2$.

Theorem 4.6.6 *Let $(\mathbf{M}, \eta, [A])$ be a semi-Randers space and η the Minkowski metric. Let us assume that*

1. *The auto-parallel curves of unit velocity of the connections ${}^L\nabla$ and $\langle {}^L\nabla \rangle$ are defined for time t , the time coordinate measured in the laboratory frame $Z = \frac{\partial}{\partial t}$.*
2. *The ultra-relativistic limit holds: $E(x) \gg 1$ for all $x \in \mathbf{K}$.*
3. *The distribution function is narrow in the sense that $\alpha \ll 1$ for all $x \in \mathbf{K}$ in the laboratory frame.*
4. *The following inequality holds,*

$$|\theta^2(t) - \bar{\theta}^2(t)| \ll 1,$$

5. *The support of the distribution function f is invariant under the flow of the Lorentz force equation.*
6. *The change in the energy function is adiabatic: $\frac{d}{dt} \log E \ll 1$.*

Then for the same arbitrary initial condition $(x(0), \dot{x}(0))$, the solutions of the equations

$${}^L\nabla_{\dot{x}} \dot{x} = 0, \quad \langle {}^L\nabla \rangle_{\dot{x}} \dot{x} = 0$$

are such that

$$\|\tilde{x}(t) - x(t)\|_{\bar{\eta}_Z} \leq 2(C(x(t))\|\mathbf{F}\|_{\bar{\eta}_Z}(x(t)) + C_2^2(x(t))(1 + B_2(x(t))\alpha))\alpha^2 E^{-2}(x) t^2 + \mathcal{O}(\alpha^4), \quad (4.6.10)$$

where the functions $C(x(t))$, $C_i(x(t))$ and $B_i(x(t))$ are bounded by constants of order 1.

Proof: At the instant t , we calculate the distance measured in the laboratory frame between $x(t)$ and $\tilde{x}(t)$, solutions of the geodesic equations of the connections ${}^L\nabla$ and $\langle {}^L\nabla \rangle$ respectively. Both geodesics have the same initial conditions $(x(0), \dot{x}(0))$. Let us start writing the general expression of the solution of the Lorentz force equation for those initial conditions:

$$x^i(t) = x^i(0) + \int_0^t ds \left(\dot{x}^i(0) + \int_0^s dl \ddot{x}^i(l) \right). \quad (4.6.11)$$

Since the initial conditions for both geodesics are the same, the equivalent relation for the geodesics of the averaged connection is

$$\tilde{x}^i(t) = x^i(0) + \int_0^t ds \left(\dot{x}^i(0) + \int_0^s dl \ddot{\tilde{x}}^i(l) \right). \quad (4.6.12)$$

We estimate the *distance* between both solutions at the instant t . Since we know the distance between the connections ${}^L\nabla$ and $\pi^* \langle {}^L\nabla \rangle$, it is possible to give a natural bound for the distance between the solutions. The main tool that we use is the smoothness theorem on the dependence of solutions of differential equations on the external parameters (for instance [21, *Appendix 1*] or [40, *chapter 1*] or in the *appendix* of this thesis).

Let us consider the family of connections depending on the distance between the two connections

$$\xi_{max} = d_{\bar{\eta}}({}^L\nabla, \langle {}^L\nabla \rangle)$$

given by the convex sum:

$$\xi\nabla := \frac{1}{\xi_{max}}(\xi_{max} - \xi){}^L\nabla + \frac{1}{\xi_{max}}\xi\pi^*\langle {}^L\nabla \rangle, \quad \xi \in [0, \xi_{max}]. \quad (4.6.13)$$

For $\xi = 0$ one has $\xi\nabla = {}^L\nabla$, while for $\xi = \xi_{max}$ one has the averaged connection. Using the result of the smoothness of the solutions of the differential equations, one can expand the solution ξx^i of the geodesic equation for the connection with

parameter ξ . The second derivative with respect to the coordinate time t reads

$${}^\xi \ddot{x}^i = {}^0 \ddot{x}^i + (\partial_\xi {}^\xi \ddot{x}^i)|_{\xi=0} \cdot \xi + \mathcal{O}(\xi^2). \quad (4.6.14)$$

We need to bound the derivative $(\partial_\xi {}^\xi \ddot{x}^i)|_{\xi=0}$. The first think is that if \dot{x} (the tangent velocity vector to a Lorentz geodesic) is not on the support of the distribution f , this derivative can be done arbitrary large. From the formula (4.6.13) and (4.6.14), one obtains:

$$(\partial_\xi {}^\xi \ddot{x}^i)|_{\xi=0} = \frac{1}{\xi_{max}} \cdot ({}^L \nabla_{\dot{x}} \dot{x} - \pi^* < {}^L \nabla_{>\dot{x}} \dot{x})^i,$$

where $\dot{x}(t)$ is the solution of ${}^L \nabla_{\dot{x}} \dot{x} = 0$ with the given initial conditions. This is because the support of the distribution function f is invariant under the flow of the Lorentz force equation. We are dividing by the distance ξ_{max} , thus the derivative is such that its norm is bounded by 1,

$$\|(\partial_\xi {}^\xi \ddot{x})|_{\xi=0}\|_{\bar{\eta}} = \frac{1}{\xi_{max}} \cdot \|({}^L \nabla_{\dot{x}} \dot{x} - \pi^* < {}^L \nabla_{>\dot{x}} \dot{x})\|_{\bar{\eta}} \leq 1,$$

because of the definition of ξ_{max} and the formula (4.6.14). Note that for writing this condition, it is essential that the support of the distribution f must be invariant under the flow defined by the Lorentz force equation: if this is not the case, the parameter ξ_{max} can not be defined and the difference on the covariant derivatives cannot be bounded. On the other hand the relations between proper times and coordinate time in the laboratory frame are

$$d\tau = \gamma^{-1} dt, \quad d\tilde{\tau} = \tilde{\gamma}^{-1} dt.$$

This implies the following relation between derivatives,

$$\frac{d}{dt} = \gamma^{-1} \frac{d}{d\tau}, \quad \frac{d}{dt} = \tilde{\gamma}^{-1} \frac{d}{d\tilde{\tau}}.$$

Using the hypotheses $|\theta^2(t) - \bar{\theta}^2(t)| \ll 1$ and $\frac{d}{dt} \log E \ll 1$, one obtains the following relation,

$$\|\tilde{x}(t) - x(t)\|_{\bar{\eta}} \leq 2t \int_0^t dl E^{-2} \left\| \frac{d^2 \tilde{x}^i(l)}{dl^2} - \frac{d^2 x^i(l)}{dl^2} \right\|_{\bar{\eta}},$$

where by the adiabatic hypothesis, the time derivatives of the energy function have been dropped out; the factor 2 comes from the bound of the term which contains the derivative of the energy. \square

Corollary 4.6.7 *Let $(\mathbf{M}, \bar{\eta})$ be as before and such that there is a global bound for $\|\mathbf{F}\|_{\bar{\eta}}(x) \leq \|\mathbf{F}\|_{\bar{\eta}} < \infty$, the energy function is bounded from below by a constant E and the curves $x(t)$ and $\tilde{x}(t)$ are compact. Then there are some constants C_i such that $C_i(x) < C_i$ and the following relation holds*

$$\|\tilde{x}(t) - x(t)\|_{\bar{\eta}} \leq 2(C\|\mathbf{F}\|_{\bar{\eta}} + C_2^2(1 + \alpha))\alpha^2 E^{-2} t^2 + \mathcal{O}(\alpha^4), \quad (4.6.15)$$

with \mathbf{F} the maximal value of $\mathbf{F}(x)$ attached along the compact curves $x(t)$ and $\tilde{x}(t)$.

Remarks

1. In the above result the 2-form \mathbf{F} is physically interpreted as the Faraday form.
2. Global bounds occurs in two possible scenarios:
 - (a) When manifold \mathbf{M} is compact. In this case, one has to consider space-like boundaries, since it is well-known that if the space-time is compact without boundaries, there exists closed time-like curves ([23, pg 58]) and this violates causality.
 - (b) The trajectories that we consider are compact in space and bounded in time between an initial time $t = 0$ and a final time $t = T$. In this case, one can define an *effective compact space-time manifold* with a spurious boundary and apply the first case, with the exclusion of closed time-like curves, which do not occur in physical situations.

3. We have assumed in our calculations that the external field \mathbf{F} does not depend on the energy of the beam of particles. However, this is not necessarily the case in some situations (as in betatron accelerator machines [9,10]).
4. There are effects which could reduce the beam size (adiabatic damping and Landau Damping [9]). If this happens, there is a strong reduction of the size of the dispersion in energy and momenta of the beam. This implies that one can describe this as an effective exponent $E^{-2+\beta}$ with $\beta < 0$. Then our results are safe under these kind of effects.
5. That the curves $x(t)$ and $\tilde{x}(t)$ have compact image has a physical interpretation: all the trajectories start in a source region and finish in a target region.
6. Although we have chosen $Z = \frac{\partial}{\partial t}$ to be the observer corresponding to the laboratory frame, the same kind of calculations can be performed in any frame.

Theorem 4.6.8 *Under the same hypothesis as in theorem 4.6.6, the difference between the tangent vectors is given by*

$$\|\dot{\tilde{x}}(t) - \dot{x}(t)\|_{\bar{\eta}} \leq (K(x)\|\mathbf{F}\|_{\bar{\eta}}(x) + K_2^2(1 + D_2(x)\alpha)\alpha^2) E^{-1}t + \mathcal{O}(\alpha^4). \quad (4.6.16)$$

with $K(x)$, $K_2(x)$ and $D_2(x)$ functions bounded by constants bounded by respective constants of order 1.

Proof: The proof of this theorem is similar to the proof of *Theorem (4.6.6)*, although based on the following formula for the tangent velocity field along a curve:

$$\dot{x}(t) = \dot{x}(0) + \int_0^t \ddot{x}(l)dl. \quad (4.6.17)$$

□

Corollary 4.6.9 *Under the same hypothesis than in Corollary 4.6.7, there are some constants of order 1, K , K_2 and D_2 , such that $K(x) \leq K$, $K(x)_2 \leq K_2$, $D_2(x) \leq D_2$*

and the following relation holds,

$$\|\dot{\hat{x}}(t) - \dot{x}(t)\|_{\bar{\eta}} \leq (K\|\mathbf{F}\|_{\bar{\eta}}(x) + K_2^2(1 + D_2(x)\alpha)\alpha^2) E^{-1} t + \mathcal{O}(\alpha^4). \quad (4.6.18)$$

4.7 Discussion

4.7.1 Structural stability of the approximation

$${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$$

In the proof of proposition (4.6.3) there is a cancelation of the leading orders of transversal and longitudinal contributions when we calculate the difference between $< {}^L\nabla >$ and ${}^L\nabla$. The contribution coming from the *transversal* terms was:

$$\begin{aligned} \frac{1}{2}\mathbf{F}^i{}_m(x)(\delta^m(x, y)\eta_{jk} - \eta_{js}\eta_{kl}\delta^{msj}(x, y)y^jy^k) &= \frac{1}{2}\mathbf{F}^i{}_m(x)(\delta^m(x, y) - \delta^m(x, y) \\ -2 < \hat{y}^m > (x)\delta^s(x, y)\eta_{sj}y^j) - \mathcal{O}_2^m(\delta^2) - \mathcal{O}_3^m(\delta^3)\mathbf{F}^i{}_m(x) &= \\ = -\mathbf{F}^i{}_m(x)(< \hat{y}^m > (x)\delta^s(x, y)\eta_{sj}y^j) - \mathcal{O}_2^m(\delta^2) - \mathcal{O}_3^m(\delta^3)\mathbf{F}^i{}_m(x). \end{aligned}$$

The contribution coming from the *longitudinal* terms was

$$\frac{1}{2}(\mathbf{F}^i{}_j(x)(< \hat{y}^m > (x) - y^m)\eta_{mk} + \mathbf{F}^i{}_k(x)(< \hat{y}^m > (x) - y^m)\eta_{mj})y^jy^k = \mathbf{F}^i{}_j(x)y^j\delta^k(x, y)y_k.$$

The reason for this cancelation of the first order term in δ is based on the formal structure of the connection ${}^L\nabla$. This structure has a two-fold origin:

1. The definition we have adopted for a general non-linear connection and
2. The structure of the Lorentz force equation.

The cancelation is independent of the details of the distribution, even if α is not small. It is independent of the value of \mathbf{F} .

The cancelation of the linear terms can be written in the following way:

$${}^L\nabla_y y - \pi^* < {}^L\nabla >_y y = \mathcal{O}(\alpha^2(x)).$$

Let us consider the connection

$$\tilde{\nabla} = {}^L\nabla - T, \quad T^i{}_{jk} = \mathbf{F}^i{}_m(x) \frac{y^m}{2\sqrt{\eta(y,y)}} \left(\eta_{jk} - \frac{1}{\eta(y,y)} \eta_{js} \eta_{kt} y^s y^t \right).$$

The connection $\tilde{\nabla}$ is such that its auto-parallel curves coincide with the solutions of the Lorentz force equation. It is also gauge invariant. Therefore, it is also a *good* candidate for a geometrization of the Lorentz force equation. However, if we calculate the analogous difference $\tilde{\nabla}_y y - \pi^* \langle \tilde{\nabla} \rangle_y y$ we get that in general it is linear in δ , which implies

$$\tilde{\nabla}_y y - \pi^* \langle \tilde{\nabla} \rangle_y y = \mathcal{O}(\alpha(x)).$$

Finally, if we consider the covariant derivative ${}^x\nabla$ associated with a spray vector field $\chi \in \Gamma\mathbf{TN}$, we obtain

$${}^x\nabla_y y - \pi^* \langle {}^x\nabla \rangle = \mathcal{O}(\alpha(x)).$$

We can write this relations as

$$\lim_{\alpha \rightarrow 0} \frac{{}^L\nabla_y y - \pi^* \langle {}^L\nabla \rangle_y y}{\alpha} = 0, \quad (4.7.1)$$

$$\lim_{\alpha \rightarrow 0} \frac{\tilde{\nabla}_y y - \pi^* \langle \tilde{\nabla} \rangle_y y}{\alpha} \neq 0, \quad (4.7.2)$$

$$\lim_{\alpha \rightarrow 0} \frac{{}^x\nabla_y y - \pi^* \langle {}^x\nabla \rangle_y y}{\alpha} \neq 0. \quad (4.7.3)$$

This fact can be stated in the following way. Let us denote a 1-parameter family of linear connections by $\nabla(\alpha)$ such that both $\nabla(\alpha)$ and $\nabla(\alpha + h)$ are defined. Then one can define the derivative operator

$$\lim_{h \rightarrow 0} \frac{\nabla(\alpha + h)_y y - \nabla(\alpha)_y y}{h},$$

which is formally the differential (in the way we define, it is the Gateaux differential [52, *chapter 5*]) of the operator

$$\nabla(\alpha, y) : \mathbf{I} \longrightarrow \mathbf{T}_x \mathbf{M}$$

$$\alpha \mapsto {}^L \nabla(\alpha)_y y.$$

for a fixed $y \in \text{supp}(f_x)$ and such that $\alpha, \alpha + h \in \mathbf{I}$. The notation makes sense such that for each y there is a given operator $\nabla_y(\alpha)y$. Then equations (4.7.1), (4.7.2) and (4.7.3) can be stated in terms of derivatives. In particular, since $\nabla(\alpha = 0, y) = {}^L \nabla_y y$, one has that equation (4.7.1) is equivalent to the statement that ${}^L \nabla_y y$ is a critical value of $\nabla(\alpha, y)$. Therefore one can write

$$\pi^* < {}^L \nabla_y y > = \pi^* < \nabla(0, y) > + \frac{\alpha^2}{2} \frac{d^2}{d\alpha^2} \Big|_{\alpha=0} \nabla(\alpha, y) + \mathcal{O}(\alpha^3).$$

When the support of the distribution is invariant under the flow of the Lorentz equation, for $\alpha = 0$ one has that ${}^L \nabla = \pi^* < {}^L \nabla >$. Therefore,

$$< {}^L \nabla_y y > = {}^L \nabla_y y + \frac{\alpha^2}{2} \frac{d^2}{d\alpha^2} \Big|_{\alpha=0} \nabla(\alpha, y) + \mathcal{O}(\alpha^3). \quad (4.7.4)$$

for any $y \in \Sigma$. For arbitrary distribution functions $\frac{\alpha^2}{2} \frac{d^2}{d\alpha} \Big|_{\alpha=0} \nabla(\alpha, y) \neq 0$.

On the other hand, equation (4.7.2) can be rewritten, following the same steps

$$\pi^* < \tilde{\nabla}_y y > = \tilde{\nabla}_y y + \alpha \frac{d}{d\alpha} \Big|_{\alpha=0} \tilde{\nabla}(\alpha, y) + \mathcal{O}(\alpha^2) \quad (4.7.5)$$

for any $y \in \Sigma$. For arbitrary distribution functions $\frac{d}{d\alpha} \Big|_{\alpha=0} \tilde{\nabla}(\alpha, y) \neq 0$. Also note that $0 = {}^L \nabla_y y = \tilde{\nabla}_y y$ for any $y \in \Sigma$, which coincides with the Lorentz force equation.

In the case of a connection ${}^x \nabla$ obtained from an arbitrary spray χ , the relation is

$$\pi^* < {}^x \nabla_y y > = {}^x \nabla_y y + \alpha \frac{d}{d\alpha} \Big|_{\alpha=0} {}^x \nabla(\alpha, y) + \mathcal{O}(\alpha^2). \quad (4.7.6)$$

Apart from the relevance for calculational purposes, it is interesting to know if there is some reason why (4.7.4) holds for the Lorentz connection ${}^L\nabla$ obtained from the Berwald connection associated to the spray ${}^L\chi$. If one changes the way we obtain the connection from the spray, or changes the spray (for instance using the connection $\tilde{\nabla}$ instead of ${}^L\nabla$), one obtains conditions of the type (4.7.5). If one performs a similar calculation for a general connection, a differential equation like (4.7.6) is obtained. This suggests that there are two factors in obtaining the relation (4.7.4):

1. The choice of the non-linear connection as the Berwald-type connection applied to the Lorentz spray ${}^L\chi$.
2. The particular structure of the Lorentz equation. This is apparent in the calculations in the proof of *proposition* (4.6.3).

The conclusion is that the non-linear Berwald connection obtained from the Lorentz force equation is structurally stable with respect to the parameter α . As compared with other connections, this property happens (maybe only) for the Berwald-type connection. This notion of stability is similar to the one presented in [61, *chapter 3*].

4.7.2 On the hypotheses on which the approximation

${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$ is based

We would like to discuss some reflections and interpretations of the hypotheses of *theorem* (4.6.6) and subsequent results. Some of the hypotheses are not essential to perform the approximation, but are very useful in the calculations and in writing the asymptotic expressions. Therefore, we can differentiate between fundamental hypotheses, which are

1. The auto-parallel curves of unit velocity of the connections ${}^L\nabla$ and $< {}^L\nabla >$ are defined for the time t , which is the coordinate time measured in the laboratory frame. This is an hypothesis on existence, since nothing can be done if the curves are not defined for the parameter that we are speaking about.

2. The support of the distribution function f is invariant under the flow of the Lorentz force equation. If this hypothesis does not hold, the relation between the averaged connection and the Lorentz connection is arbitrary. Therefore, this is a fundamental hypothesis. Note that one does not need to assume that f is a solution of the Vlasov equation in the sense that the condition is weaker. This implies that our result holds for alternative kinetic models for the distribution function f .
3. The dynamics occurs in the ultra-relativistic limit, $E(x) \gg 1$ for all $x \in \mathbf{M}$. This is a hypothesis which in principle is not fundamental for the calculation but it is fundamental for the approximation ${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$ to be good. If this hypothesis does not hold, the difference between the solutions of the two differential equations will not be as small as in the ultra-relativistic limit and the expressions will be more involved.
4. The distribution function is narrow, $\alpha \ll 1$ for all $x \in \mathbf{M}$. This is useful to interpret the formulas as asymptotic series in α . Note that for current accelerators, this condition holds.

Apart from discussed above, there are other hypotheses which are not fundamental to the results, although they are helpful in the calculations

1. The following inequality holds

$$|\theta^2 - \bar{\theta}^2| \ll 1.$$

This hypothesis is only used in the proof to simplify some expressions. Therefore, it is not fundamental. The approximation ${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$ can be good even if $|\theta^2 - \bar{\theta}^2| \ll 1$ is non true, but the asymptotic expressions will be more involved.

2. The adiabatic hypothesis $\frac{d \log E}{dt} \ll 1$ is not necessary for the approximation ${}^L\nabla \longrightarrow \pi^* < {}^L\nabla >$ to be good. However, this hypothesis simplifies the

calculations and the final expression. Note that this is a condition which is satisfied in actual accelerator machines.

There are limitations on the validity of the results that we have obtained in this chapter.

1. The 2-form \mathbf{F} is interpreted physically as the Faraday form of an external electromagnetic field. We have not commented on the dependence on the strength of the electromagnetic field. It is clear that for any finite α , if the electromagnetic field is too strong, the approximation will not be good. However, for delta distribution functions with invariant support by the flow of the Lorentz vector field L_χ , the approximation is always valid.
2. The main results of this chapter (*theorems* (4.6.6) and (4.6.8) and *corollaries* (4.6.7) and (4.6.9)) are not Lorentz covariant. However, as we have said before, given an arbitrary observer defined by a time-like vector field Z , it is possible to obtain similar results.

4.7.3 Range of applicability of the approximation

Let us discuss the limits of applicability of the averaged dynamics and in particular of the formula (4.6.10). We will make the assumption that during the time of evolution the γ factors are increasing. Therefore, equation (4.6.10) takes the form

$$\|\tilde{x}(t) - x(t)\|_{\bar{\eta}} \leq C \alpha^2 E^{-2}(t_0) \|\mathbf{F}\|_{\bar{\eta}} t^2.$$

Note that we are assuming $E(t) \geq E(t_0)$.

Let us assume a natural maximal spatial distance L_{max} between points on $supp(f)$. For instance, in an accelerator machine, L_{max} can be related with the diameter of the pipe: L_{max} must be smaller than it. Assume that the averaged model is a good approximation if the distance $\|\tilde{x}^i(t) - x^i(t)\|_{\bar{\eta}}$ is less than L_{max} . This can start to happen after a time evolution such that the difference $\|\tilde{x}^i({}^1t_{max}) - x^i({}^1t_{max})\|_{\bar{\eta}} =$

L_{max} . The characteristic time where the averaged model loses validity is

$${}^1t_{max} \sim \left(\frac{L_{max}}{C_1}\right)^{\frac{1}{2}} \cdot \frac{E(t_0)}{\alpha} \cdot \left(\frac{1}{\|\mathbf{F}\|}\right)^{\frac{1}{2}}. \quad (4.7.7)$$

There is a second constraint coming from the spread in velocities. Let us consider the relations (4.6.16) and (4.6.17). For any distribution function, one can expect that the approximation ${}^L\nabla \longrightarrow \langle {}^L\nabla \rangle$ to be valid until $\|\dot{\hat{x}}(t) - \dot{x}(t)\|_{\bar{\eta}}$ is of order α . This is because the unit hyperboloid is strictly convex subset of the tangent space. One writes the condition

$$\alpha = \left(K\|\mathbf{F}\|_{\bar{\eta}_Z}(x) + K_2^2(1 + D_2\alpha)\alpha^2\right) E^{-1}(t_0) {}^0t_{max} = K\|\mathbf{F}\|_{\bar{\eta}_Z}\alpha^2 E^{-1}(t_0) {}^2t_{max},$$

The maximal time calculated in this way is

$${}^2t_{max} = K \cdot \frac{E(t_0)}{\alpha} \cdot \frac{1}{\|\mathbf{F}\|_{\bar{\eta}}}. \quad (4.7.8)$$

The asymptotic behavior of the time where the approximation is broken is universal. We have to point out that the approximation can be broken before t_{max} . However,

Proposition 4.7.1 *The following consequences are true*

1. $\lim_{E \rightarrow \infty} t_{max} = \infty$, if all the other parameters are finite,
2. $\lim_{\alpha \rightarrow 0} t_{max} = \infty$, if all the other parameters are finite,
3. $\lim_{\|\mathbf{F}\|_{\bar{\eta}} \rightarrow 0} t_{max} = \infty$, if all the other parameters are finite.

The existence of two maximal times up to where the approximation is valid provides a definition of maximal length L_{max} : it is the length for which

$$\left(\frac{L_{max}}{C}\right)^{\frac{1}{2}} \cdot \frac{E(t_0)}{\alpha} \cdot \left(\frac{1}{\|\mathbf{F}\|_{\bar{\eta}}}\right)^{\frac{1}{2}} = K \cdot \frac{E(t_0)}{\alpha} \cdot \frac{1}{\|\mathbf{F}\|_{\bar{\eta}}}.$$

All the quantities appearing are known, therefore one has

$$L_{max} = A \cdot \frac{1}{\|\mathbf{F}\|_{\bar{\eta}}}, \quad (4.7.9)$$

with A a constant which does not depend on x , \mathbf{F} , α or E .

If we compare this definition with $\bar{L}_{max} := \text{diam}(\pi(\text{supp}(f)))$, we have a criterion for a definition of *weak* 2-form:

Definition 4.7.2 *The 2-form \mathbf{F} is said to be weak iff $L_{max} \leq \bar{L}_{max}$.*

Proposition 4.7.3 *If the 2-form \mathbf{F} is not weak, then the approximation ${}^L\nabla \rightarrow < {}^L\nabla >$ is not valid.*

Proof: If $\bar{L}_{max} \geq L_{max}$ the system cannot reach the boundary $\partial(\pi(\text{supp}(f)))$ before the approximation breaks down because the 2-form \mathbf{F} produces undesirable effects on the dynamics of the system. \square

Finally, we have the following *corollary*:

Corollary 4.7.4 *Under the same hypothesis of theorem 4.6.6, the following hypothesis hold:*

1. *In the limit $\alpha \rightarrow 0$ the Lorentz force equation and the averaged Lorentz force equation coincide.*
2. *In the limit $E \rightarrow \infty$ the Lorentz force equation and the averaged Lorentz force equation coincide.*

Chapter 5

Charged cold fluid model from the Vlasov model

5.1 Introduction

Despite limitations concerning the mathematical description of the discrete nature of the particles comprising a plasma, modeling the dynamics of relativistic non-neutral plasmas and charged particle beams by fluid models is common place. The relative simplicity of these models, compared with the corresponding kinetic models makes them appealing.

We propose in this *chapter* another justification for the use of fluid models in beam dynamics. We will concentrate on the charged cold fluid model. However, we should notice that the same philosophy is also applicable to more sophisticated models.

In high intensity beam accelerator machines, each bunch of a beam contains a large number of identical particles contained in a small phase-space region. In such conditions, a number of the order $10^9 - 10^{11}$ charged particles move *together* under the action of both external and internal electromagnetic fields. Often in modern applications, such bunches of particles move ultra-relativistically.

One is interested in modeling these physical systems in such a way that:

1. The model for a bunch of particles must be *simple*, in order to be useful in numerical simulations of beam dynamics and for analytical treatment,
2. It allows for stability analysis and a qualitative understanding of the dynamical behavior of the system. Three dimensional numerical simulations can be also desirable.

The standard approach has been to use fluid models as an approximation to a kinetic model. These derivations of fluid models from kinetic models can be found for instance in [8, 37-39] and references therein. They are based on some assumptions, usually in the form of equations of state for fluids or assumptions on the higher moments of the distribution function f . These constraints are necessary in order to close the hierarchy of moments of the distribution function and to have a sufficient number of differential relations to determine the remaining moments. This is a general feature of all the derivations of fluid models from kinetic theory: a truncation scheme is required for the fluid model to be predictive.

We present in this *chapter* a new *justification* of the charged cold fluid model from the framework of kinetic theory. The novelty of the new approach is that it uses *natural hypotheses* suitable for particle accelerator machines and exploits only the mathematical structure of the classical electrodynamics of charged point particles interacting with external electromagnetic fields. We estimate the covariant derivative of the mean velocity calculated with the one-particle distribution function. This is given as an asymptotic formula in terms of the time of the evolution, diameter of the distribution and the energy of the beam. The charged cold fluid model is described by only one *dynamical* variable, the normalized mean velocity field. The variance and the heat flow tensor are not necessarily zero, but are finite and given externally. In our treatment both the fluid energy tensor and the flux tensor are assumed to be given. Our aim is not to give an equation for the mean velocity field, but to evaluate how much certain differential expressions (formally equivalent to the charged cold fluid model equations) differ from zero. Then one can stipulate the validity of the model from the estimates of the corresponding differential expressions. On the other hand, in the models presented for instance in [8, 37-39], the variance and

the covariant heat flow are dynamical variables and a system of partial differential equations is used to determine the dynamics of these fields. We think the analysis performed in this *chapter*, can be extended to other fluids models in future work.

The method used in this thesis to obtain these results is the following:

1. One considers the results from *chapter 4*, which compares the solutions of the Lorentz connection with the solutions of the averaged Lorentz force equation. In this sense, we are under the same hypotheses as in *theorems 4.6.6* and *4.6.8*. We will use the bounds of the differences of the corresponding geodesics.
2. It happens that under the same assumptions as used for the particle dynamics, the corresponding solutions of the Vlasov equation f and the averaged Vlasov equation \tilde{f} are similar. This result is based on the comparison results of the point particle dynamics.
3. Each of the distributions f and \tilde{f} determine a different mean velocity field. One can prove under the same hypotheses that these mean velocity vector fields are similar. This means that the difference between them is controlled by powers of small parameters.
4. Finally, we show that the auto-parallel condition of the velocity field of the averaged Vlasov equation associated with the averaged dynamics is controlled by the diameter of the distribution f . Together with the above point, this result allows us to provide estimates for the auto-parallel condition of the mean velocity field of the solution of the Vlasov equation.

Therefore, the methods presented here and the usual derivations of the fluid models contained in [8, 37-39] are different. The standard approaches assume an asymptotic expansion of the differential equations for the moments, in terms of a perturbation parameter which is similar to the diameter α of the distribution function. These asymptotic expansions are particularized at low orders as a *truncation scheme* in the hierarchy of moments. Those references discuss systems of partial differential equations which are self-contained and consistent with physical constraints and with the

asymptotic expansions. On the other hand, our approach is based on the structure of the Lorentz force equation of a charged point particle, which lies at the basis of the kinetic models. Written in a geometric way, the Lorentz force equation is replaced by the averaged Lorentz force equation. The key point is that the averaged Lorentz connection admits normal coordinates, which simplifies in a fundamental way our calculations. Then under some regularity assumptions on the distribution function, we can place bounds on the differential expression of interest.

We have assumed that the distribution functions are smooth (at least of class \mathcal{C}^1) in the coordinates x^i . Although we do not currently have a proof that we can extend our results to bigger functional spaces for the distribution functions, since the main results are written in terms of Sobolev norms, it is conjectured that they can be extended to Sobolev spaces. Indeed our proofs suggest that we require smoothness on the x coordinates and the existence of weak derivatives on the y coordinates.

5.2 Comparison of the solutions of the Vlasov and averaged Vlasov equations

In this *section* we estimate the difference between the solutions of the Liouville equations associated with the averaged Lorentz connection and the original Lorentz connection. The Liouville equation associated to the Lorentz force equation is called Vlasov's equation in kinetic theory. In a similar way, we call to the Liouville equation associated with the averaged Lorentz force equation the averaged Vlasov equation.

5.2.1 Examples of Liouville equations

Given a non-linear connection characterized by the *second order* vector field $\chi \in \mathbf{TTN}$, the associated Liouville equation is $\chi(f) = 0$. In the following two examples presented here, which are related with our purposes.

1. From the coefficients of the Lorentz connection ${}^L\Gamma^i_{jk}(x, y)$ one can recover the spray coefficients ${}^LG^i(x, y)$, using the homogeneous properties on y of ${}^LG^i(x, y)$

and Euler's theorem of positive homogeneous functions. In particular, the spray coefficients are

$$\begin{aligned}
{}^L G^i(x, y) &= {}^L \Gamma^i_{jk}(x, y) y^j y^k = \left(\eta \Gamma^i_{jk} + \frac{1}{2\sqrt{\eta(y, y)}} (\mathbf{F}^i_j(x) y^m \eta_{mk} + \right. \\
&\quad \left. + \mathbf{F}^i_k(x) y^m \eta_{mj}) + \mathbf{F}^i_m(x) \frac{y^m}{2\sqrt{\eta(y, y)}} (\eta_{jk} - \frac{1}{\eta(y, y)} \eta_{js} \eta_{kl} y^s y^l) \right) y^j y^k = \\
&= \left(\eta \Gamma^i_{jk} + \frac{1}{2\sqrt{\eta(y, y)}} (\mathbf{F}^i_j(x) y^m \eta_{mk} + \mathbf{F}^i_k(x) y^m \eta_{mj}) \right) y^j y^k = \\
&= \eta \Gamma^i_{jk} y^j y^k + \sqrt{\eta(y, y)} \mathbf{F}^i_j(x) y^j.
\end{aligned}$$

Then one can define the vector field ${}^L \chi$:

$${}^L \chi(x, y) = y^i \frac{\partial}{\partial x^i} - \left(\eta \Gamma^i_{jk}(x) y^j y^k + \sqrt{\eta(y, y)} \mathbf{F}^i_j(x) y^j \right) \frac{\partial}{\partial y^i}.$$

2. A similar procedure applies to the averaged Lorentz Vlasov vector field. In this case, however, we do not have the simplifications given above. Therefore, the spray coefficients are

$$\begin{aligned}
< {}^L G^i > (x, y) &= {}^L \Gamma^i_{jk}(x, y) y^j y^k = \left(\eta \Gamma^i_{jk}(x) + < \frac{1}{2\sqrt{\eta(y, y)}} (\mathbf{F}^i_j(x) y^m \eta_{mk} + \right. \\
&\quad \left. + \mathbf{F}^i_k(x) y^m \eta_{mj}) + \mathbf{F}^i_m(x) \frac{y^m}{2\sqrt{\eta(y, y)}} (\eta_{jk} - \frac{1}{\eta(y, y)} \eta_{js} \eta_{kl} y^s y^l) > \right) y^j y^k.
\end{aligned}$$

If $y \in \Sigma$, semi-spray coefficients can be simplified to

$$\begin{aligned}
< {}^L G^i > (x, y)|_{\Sigma} &= \left(\eta \Gamma^i_{jk}(x) + < \frac{1}{2} (\mathbf{F}^i_j(x) y^m \eta_{mk} + \mathbf{F}^i_k(x) y^m \eta_{mj}) + \right. \\
&\quad \left. + \mathbf{F}^i_m(x) \frac{y^m}{2} (\eta_{jk} - \eta_{js} \eta_{kl} y^s y^l) > \right) y^j y^k = \\
&= \left(\eta \Gamma^i_{jk}(x) + \frac{1}{2} (\mathbf{F}^i_j(x) < y^m > \eta_{mk} + \mathbf{F}^i_k(x) < y^m > \eta_{mj}) + \right. \\
&\quad \left. + \mathbf{F}^i_m(x) (\eta_{jk} < \frac{y^m}{2} > - \eta_{js} \eta_{kl} < \frac{y^m}{2} y^s y^l >) \right) y^j y^k.
\end{aligned}$$

The averaged Vlasov vector field can be written in a similar way as before,

$$\begin{aligned} \langle {}^L\chi \rangle|_{\Sigma} = & y^i \frac{\partial}{\partial x^i} - \left(\eta \Gamma^i{}_{jk} y^j y^k + \frac{1}{2} (\mathbf{F}^i{}_j(x) \langle y^m \rangle \eta_{mk} + \mathbf{F}^i{}_k(x) \langle y^m \rangle \eta_{mj}) + \right. \\ & \left. + \mathbf{F}^i{}_m(x) (\eta_{jk} \langle \frac{y^m}{2} \rangle - \eta_{js} \eta_{kl} \langle \frac{y^m}{2} y^s y^l \rangle) \right) y^j y^k \frac{\partial}{\partial y^i}. \end{aligned}$$

5.2.2 Comparison between the solutions of the Vlasov equation and the averaged Vlasov equation

In the following (\mathbf{M}, η) is Minkowski space, since we will use *theorems (4.6.6)* and *theorem (4.6.8)*. The Riemannian metric $\bar{\eta}_Z$ is determined by the vector $Z = \frac{\partial}{\partial t}$. Since we will use the Euclidean metric associated with $Z = \frac{\partial}{\partial t}$, we simplify the notation and employ $\bar{\eta}$ in place of $\bar{\eta}_Z$. $Z = \frac{\partial}{\partial t}$ will be the observer that we call laboratory frame. We will restrict our attention to a compact domain $\mathbf{K} \subset \mathbf{M}$.

Proposition 5.2.1 *Let f and \tilde{f} be solutions of the Vlasov equation ${}^L\chi(f) = 0$ and the averaged Vlasov equation $\langle {}^L\chi \rangle(\tilde{f}) = 0$, where ${}^L\chi$ and $\langle {}^L\chi \rangle$ are the spray vector fields obtained from the non-linear connections ${}^L\nabla$ and $\langle {}^L\nabla \rangle$. Let us assume the same hypotheses as those in theorem 4.6.6. Then for the solutions of the Vlasov and averaged Vlasov's equation with the same initial conditions, one has the relation*

$$\begin{aligned} |f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))| &< (\tilde{C}(x) \|\mathbf{F}\|_{\bar{\eta}} C_2^2(x) (1 + B_2(x)\alpha)) \alpha^2 E^{-2} t^2 + \\ &+ (\tilde{K}(x) \|\mathbf{F}\|_{\bar{\eta}}(x) K_2^2(1 + D_2(x)\alpha)) \alpha^2 E^{-1} t \end{aligned} \quad (5.2.1)$$

for some functions $\tilde{C}(x(t))$ $\tilde{K}(x(t))$ along the geodesic of the Lorentz connection.

Proof: f and \tilde{f} are solutions of the corresponding Vlasov and averaged Vlasov equations respectively. Therefore, f and \tilde{f} are constant along the corresponding auto-parallel curves; $x(t)$ and $\tilde{x}(t)$ are the projections on the space-time manifold \mathbf{M} of the integral curves of the vector fields ${}^L\chi$ and $\langle {}^L\chi \rangle$. Here t is the time-parameter in the laboratory frame determined by the vector field $\frac{d}{dt}$. Then the

Vlasov and averaged Vlasov equation can be written as

$${}^L\chi f = \frac{d}{dt}f(x(t), \dot{x}(t)) = 0, \quad < {}^L\chi > \tilde{f} = \frac{d}{dt}\tilde{f}(\tilde{x}(t), \dot{\tilde{x}}(t)) = 0.$$

For the same initial conditions, the geodesic curves corresponding to the connections ${}^L\nabla$ and $< {}^L\nabla >$ are nearby curves at the instant t in the way described by *theorem 4.6.6*.

Let us introduce the family of *interpolating connections*,

$${}^L\nabla_\epsilon := (1 - \epsilon) {}^L\nabla + \epsilon < {}^L\nabla >, \quad \epsilon \in [0, 1].$$

Each of them has an associated spray vector field ${}^L\chi_\epsilon$. Therefore, let us consider $f_\epsilon(x, y)$ to be the solution of the following Liouville equation ${}^L\chi_\epsilon f_\epsilon = 0$ for some given initial conditions. Since the dependence on (ϵ, x, y) of the vector field ${}^L\chi_\epsilon$ is \mathcal{C}^1 , the solutions of the Liouville equation are Lipschitz with respect to the parameter ϵ . We can see this fact in the following way. The Liouville equation can be written as

$${}^L\chi_\epsilon f_\epsilon = 0 \Leftrightarrow \frac{d}{dt}f(x_\epsilon(t), y_\epsilon(t)) = 0,$$

where $(x_\epsilon(t), y_\epsilon(t))$ is an integral curve of the vector field ${}^L\chi_\epsilon$ restricted to the unit hyperboloid bundle and such that it is parameterized by the coordinate time t . Then one can use standard results from the theory of ordinary differential equations to study the smoothness properties of the solutions of the above equation. In particular, the connection coefficients for the interpolating connection are,

$$({}^L\Gamma_\epsilon)^i{}_{jk} = (1 - \epsilon) {}^L\Gamma^i{}_{jk} + \epsilon < {}^L\Gamma^i{}_{jk} > .$$

From the formula (4.5.1) for the coefficients ${}^L\Gamma^i{}_{jk}(x, y)$ one can check that $({}^L\Gamma_\epsilon)^i{}_{jk}$ are smooth functions in an open set of time-like vectors y and the parameter ϵ . From here it follows the Lipschitz condition for f_ϵ in ϵ .

We will give an upper bound for the difference $|f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))|$. Note that in this expression the point where both f and \tilde{f} are evaluated are $(t, x(t), \tilde{x}(t))$.

In order to achieve this, standard results on the smoothness of the solution of differential equations are used (see *chapter 1* of [40]). In particular we use that for each $(\bar{\epsilon}, \bar{x}(s), \bar{y}(s))$, there is an open neighborhood $\mathbf{U}_{\bar{\epsilon}}$ of $[0, 1] \times \text{supp}(f)_{\mathbf{TK}}$ containing $(\bar{\epsilon}, \bar{x}(t), \bar{y}(t))$ such that the solutions of the differential equations are Lipschitz in $\mathbf{U}_{\bar{\epsilon}}$. Therefore, using the Lipschitz condition, one obtains the bound

$$\begin{aligned} |f^\epsilon(t, x_\epsilon(t), \dot{x}_\epsilon(t)) - f^{\bar{\epsilon}}(t, x_{\bar{\epsilon}}(t), \dot{x}_{\bar{\epsilon}}(t))| &\leq c_1(\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t))\delta((\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t))) + \\ &+ c_2(\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t)) \|x_\epsilon(t) - x_{\bar{\epsilon}}(t)\|_{\bar{\eta}} + c_3(\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t)) \|\dot{x}_\epsilon(t) - \dot{x}_{\bar{\epsilon}}(t)\|_{\bar{\eta}}. \end{aligned}$$

$c_i(\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t))$ are constants which depend on the open neighborhood $\mathbf{U}_{\bar{\epsilon}}$; $\delta((\bar{\epsilon}, \bar{x}(t), \dot{\bar{x}}(t)))$ is the diameter on the ϵ component where we are applying the Lipschitz condition.

One can always choose a refinement of an open cover of $[0, 1] \times \text{supp}(f)_{\mathbf{TK}}$ such that both the Lipschitz condition, *theorem* (4.6.6) and *theorem* (4.6.8) can be applied simultaneously. Since $[0, 1]$ is compact, we can consider a finite open covering of $[0, 1]$ for each instant t . Then using the above local bound in each of the open sets \mathbf{U}_ϵ , one obtains the global bound

$$|f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))| < c_1 + c_2 \|x(t) - \tilde{x}(t)\|_{\bar{\eta}} + c_3 \|\dot{x}(t) - \dot{\tilde{x}}(t)\|_{\bar{\eta}}.$$

The constants c_i are finite (by definition of Lipschitz and by compactness of the interval $[0, 1]$). The functions f and \tilde{f} are constant along the respective geodesics. Therefore,

$$|f(t, x(t), \dot{x}(t)) - \tilde{f}(t, \tilde{x}(t), \dot{\tilde{x}}(t))| = |f(0, x(0), \dot{x}(0)) - \tilde{f}(0, \tilde{x}(0), \dot{\tilde{x}}(0))|.$$

Let us assume the same initial conditions $x(0) = \tilde{x}(0)$ and $\dot{x}(0) = \dot{\tilde{x}}(0)$ for the geodesics of the Lorentz connection. Since the difference $|f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))|$ is a smooth function of $\|x(t) - \tilde{x}(t)\|_{\bar{\eta}}$ and $\|\dot{x}(t) - \dot{\tilde{x}}(t)\|_{\bar{\eta}}$, one obtains

$$0 \leq c_1 \leq \bar{K}_1 \|x(t) - \tilde{x}(t)\|_{\bar{\eta}} + \bar{K}_1 \|\dot{x}(t) - \dot{\tilde{x}}(t)\|_{\bar{\eta}}$$

for some constants K_i . Then we have,

$$\begin{aligned} & |f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))| \leq \\ & \leq |f(t, x(t), \dot{x}(t)) - \tilde{f}(t, \tilde{x}(t), \dot{\tilde{x}}(t))| + |\tilde{f}(t, x(t), \dot{x}(t)) - \tilde{f}(t, \tilde{x}(t), \dot{\tilde{x}}(t))|. \end{aligned}$$

The first term is bounded by c_1 , which is bounded by $\bar{K}_1 \|x(t) - \tilde{x}(t)\|_{\bar{\eta}} + \bar{K}_1 \|\dot{x}(t) - \dot{\tilde{x}}(t)\|_{\bar{\eta}}$. The second term can be developed in Taylor series in the differences $\|x(t) - \tilde{x}(t)\|_{\bar{\eta}}$ and $\|\dot{x}(t) - \dot{\tilde{x}}(t)\|_{\bar{\eta}}$, since \tilde{f} is smooth. Therefore,

$$\begin{aligned} |f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))| & \leq (\tilde{C}(x(t)) \|\mathbf{F}\|_{\bar{\eta}}(x(t)) C_2^2 (1 + B_2(x(t))\alpha)) \alpha^2 E^{-2} t^2 + \\ & + (\tilde{K}(x(t)) \|\mathbf{F}\|_{\bar{\eta}}(x(t)) K_2^2 (1 + D_2(x(t))\alpha)) \alpha^2 E^{-1} t. \end{aligned}$$

□

5.3 The charged cold fluid model from the averaged Vlasov model

In the following results (\mathbf{M}, η) is Minkowski space, since we will use *theorem (4.6.6)* and *theorem (4.6.8)*. The Riemannian metric $\bar{\eta}_Z$ is determined by the vector $Z = \frac{d}{dt}$.

There is another local observer related with the vector field $\langle y \rangle$. Since the norm is not continuous on \mathbf{M} , one needs a local smoothing procedure. Given a subset of the paracompact manifold, Σ_x , we can take the induced bump function from the bump functions defined on Σ_x . Using these bump functions, we can smooth vector fields [51, pg 25].

5.3.1 Comparison of the Vlasov model with the averaged Vlasov model

Definition 5.3.1 *Given a semi-Randers space $(\mathbf{M}, \eta, [A])$, the averaged Vlasov model is defined by the dynamical variables \tilde{f} determined by*

$$\langle {}^L\chi \rangle \tilde{f} = 0, \quad (5.3.1)$$

where $\langle {}^L\chi \rangle$ is the Liouville vector field of the averaged Lorentz dynamics associated with the external electromagnetic field \mathbf{F} . The dynamical variable $f(x, y)$ defines the following

$$\tilde{V} := \int_{\Sigma_x} y \tilde{f}(x, y) d\text{vol}(x, y), \quad \text{vol}(\Sigma_x) := \int_{\Sigma_x} d\text{vol}(x, y) \tilde{f}(x, y). \quad (5.3.2)$$

Since we will use the results of *chapter 4*, the distribution function \tilde{f} is at least of type \mathcal{C}^1 in the x -coordinates and Lipschitz on the y -coordinates. Since the support $f \in f_x$ is compact, several Sobolev norms are defined [41, *chapter 3*]. We will write our results in terms of those norms.

Proposition 5.3.2 *Let $\langle {}^L\chi \rangle \tilde{f}(x, y) = 0$ and $\langle {}^L\chi \rangle f(x, y) = 0$ be such that the domain of definition of the vector field $\langle {}^L\chi \rangle$ is an open sub-manifold of Σ . Then one can reduce $\text{supp}(\tilde{f}_x) \longrightarrow \text{supp}(f_x)$ for all $x \in \mathbf{M}$.*

Proof: Let us consider the product of the functions $\tilde{f}(x, y)g(x, y)$, where $\langle {}^L\chi \rangle \tilde{f} = 0$, the function $g(x, y)$ is a *bump* function adapted to the support in \mathbf{K} of the vector field $\langle {}^L\chi \rangle$. Since both the support of $\langle {}^L\chi \rangle$ and $\text{supp}(f_x)$ are sub-sets of the paracompact manifold Σ , this function exists [32]. Therefore, we select the function $g(x, y)$ such that

$$g_x(y) = 0, \quad (x, y) \in \text{supp}(\tilde{f}_x) \setminus U(f_x), \quad g_x(y) = 0, \quad (x, y) \in \partial \text{supp}(f_x)$$

where $U(f_x) \supset \text{supp}(f_x)$ and all the derivatives are zero on $\partial \text{supp}(f_x)$. Then one

can perform the following calculation:

$$\langle {}^L\chi \rangle (\tilde{f}g) = g \langle {}^L\chi \rangle \tilde{f} + \tilde{f} \langle {}^L\nabla \rangle g = 0.$$

We can always restrict the solutions of $\langle {}^L\chi \rangle \tilde{f} = 0$ in such a way that formally $\text{supp}(\tilde{f}_x) = \text{supp}(f_x)$ and points 2 and 3 are proved. \square

Using equation (5.2.1), it follows that the error induced by the substitution $\tilde{f} \rightarrow f$ is of order α^2 . Hence, in the following calculations, when it is useful, we can use $d\text{vol}(x, y)$ as a measure and substitute $\text{supp}(\tilde{f}_x)$ by $\text{supp}(f_x)$ and \tilde{f} by f .

Let \mathbf{F} be a closed differential 2-form defining the Liouville vector field ${}^L\chi$. Let us consider the Sobolev spaces $(\mathcal{W}^{1,1}(\Sigma_x), \|\cdot\|_{1,1})$ and $(\mathcal{W}^{0,2}(\Sigma_x), \|\cdot\|_{0,2})$ [41]. Recall that the space of smooth functions is denoted by $\mathcal{F}(\Sigma_x)$ (an introduction to the notions of Sobolev spaces can be found in *appendix 5* or [40, *chapter 3*]).

Recall that we have denoted $\delta(x, y) = \langle y \rangle (x) - y$. For the next results we will restrict to the Minkowski space-time (\mathbf{M}, η) . Let us denote $\Upsilon(\text{supp}(\tilde{f}_x))$ the characteristic function of $\text{supp}(\tilde{f}_x)$.

Theorem 5.3.3 *Let \mathbf{M} be an n -dimensional space-time manifold, $\mathbf{K} \subset \mathbf{M}$ a compact domain and $\langle {}^L\chi \rangle$ the vector field associated with the averaged Lorentz force equation. Assume that:*

1. *The distribution function is such that $\tilde{f}_x, \partial_j \tilde{f}(x, \cdot) \in \mathcal{F}(\Sigma_x) \subset \mathcal{W}^{1,1}(\Sigma_x)$,*
2. *The function $\delta(x, \cdot), (\partial_j \delta)(x, \cdot) \in \mathcal{W}^{0,2}(\Sigma_x)$.*

Then

$$\|\langle {}^L\nabla \rangle_{\tilde{V}} \tilde{V}(x)\|_{\tilde{\eta}} \leq \frac{\text{vol}_E^{\frac{1}{2}}(\text{supp}(\tilde{f}_x))}{\text{vol}(\text{supp}(\tilde{f}_x))} \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) \cdot \|\tilde{f}_x\|_{1,1} \cdot \alpha^2 + O(\alpha^3), \quad (5.3.3)$$

where $\delta_x(\cdot) := \delta(x, \cdot)$ and

$$\tilde{V}^i(x) = \langle \hat{y}^i \rangle_{\tilde{f}}(x) := \frac{1}{\int_{\Sigma_x} f(x, y) d\text{vol}(x, y)} \int_{\Sigma_x} d\text{vol}(x, y) f(x, y) y^i.$$

The volumes are

$$vol(supp(\tilde{f}_x)) := vol(\Sigma_x); \quad vol_E(\Sigma_x) := \int_{\Sigma_x} \Upsilon(supp(\tilde{f}_x)) \cdot dvol(x, \tilde{y});$$

the derivative in equation (5.3.3) refers to the local frame such that the vector $U := \frac{\langle y \rangle}{\sqrt{\eta(\langle y \rangle, \langle y \rangle)}} = (U_0, \vec{0})$.

Proof: Because the averaged Lorentz connection is an affine connection on \mathbf{M} , given a point $x \in \mathbf{M}$, there is a coordinate system where the connection coefficients are zero at that point, $\langle {}^L \Gamma \rangle^i{}_{jk}(x) = 0$. Therefore, for any given point $x \in \mathbf{M}$ one can choose a *normal coordinate system* such that the *averaged Vlasov condition* holds,

$$y^j \partial_j \tilde{f}(x, y)|_x = 0. \quad (5.3.4)$$

Using this normal coordinate system, one can get a simplified expression for the covariant derivative of \tilde{V} along the integral curve of \tilde{V} :

$$\langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V} = (\tilde{V}^j \partial_j \tilde{V}^k) \frac{\partial}{\partial x^k}, \quad (5.3.5)$$

using a coordinate frame $\{\frac{\partial}{\partial x^k}, k = 0, \dots, n-1\}$. Note that this expression is not a partial differential equation because it only holds at the point x .

From the relation (5.3.5) we obtain that

$$\begin{aligned} \langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V}(x) &= \tilde{V}^i \partial_i \tilde{V}^k = \frac{1}{vol(\Sigma_x)} \int_{\Sigma_x} dvol(x, y) y^j \tilde{f}(x, y) \cdot \\ &\cdot \partial_j \left(\frac{1}{vol(\Sigma_x)} \int_{\Sigma_x} dvol(x, \hat{y}) \hat{y}^k \tilde{f}(x, \hat{y}) \right). \end{aligned}$$

It is the right hand of this equation that we shall estimate,

$$\begin{aligned} &\frac{1}{vol(supp(\tilde{f}_x))} \left(\int_{\Sigma_x} dvol(x, y) y^j \tilde{f}(x, y) \cdot \right. \\ &\left. \cdot \partial_j \left(\frac{1}{vol(\Sigma_x)} \int_{\Sigma_x} dvol(x, \hat{y}) \hat{y}^k \tilde{f}(x, \hat{y}) \right) \right) = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\text{vol}(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \left(- \frac{\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \tilde{f}(x, \hat{y}) \hat{y}^k}{\text{vol}^2(\Sigma_x)} \right. \right. \\
&\quad \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) f(x, \hat{y}) \right) \left. \right) + \frac{1}{\text{vol}(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\
&\quad \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \hat{y}^k \tilde{f}(x, \hat{y}) \right) \right) = \\
&= -\frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \tilde{f}(x, \hat{y}) \right) \langle y^k \rangle \right) + \\
&\quad + \frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\
&\quad \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \hat{y}^k \tilde{f}(x, \hat{y}) \right) \right).
\end{aligned}$$

Shifting the variable of integration $-\hat{y}^k + \langle \hat{y}^k \rangle = -\delta^k(x, \hat{y})$, one obtains the following for the above expression

$$\begin{aligned}
\langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V}(x) &= -\frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\
&\quad \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol} \hat{y} \tilde{f}(x, \hat{y}) \right) \langle y^k \rangle \right) + \\
&\quad + \frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\
&\quad \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) (\langle y^k \rangle + \delta^k(x, \hat{y})) \tilde{f}(x, \hat{y}) \right) \right) \partial_k = \\
&= \frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \delta^k(x, \hat{y}) \tilde{f}(x, \hat{y}) \right) \right) \partial_k.
\end{aligned}$$

Since $y^i \partial_i f(x, y) = 0$ at x and since \tilde{f}_x is a smooth function of y , we can Taylor expand the integrand, obtaining:

$$\begin{aligned}
&\frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\
&\quad \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \delta^k(x, \hat{y}) \left(\tilde{f}(x, y) + \frac{\partial \tilde{f}}{\partial \hat{y}^l} (\hat{y}^l - y^l) \right) \right) \right).
\end{aligned}$$

There is a coordinate system such that $\langle {}^L \Gamma \rangle^i{}_{jk}(x) = 0$ at the point x . This is

reflected in the averaged Vlasov equation, which has the form $y^j \partial_j \tilde{f}(x, y) = 0$ at one given point $x \in \mathbf{M}$. Then we get the following for the above expression

$$\left(\langle {}^L D \rangle_{\tilde{V}} \tilde{V}(x) \right)^k = \frac{1}{\text{vol}^2(\Sigma_x)} \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \\ \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \delta^k(x, \hat{y}) \frac{\partial f}{\partial \hat{y}^l} (\hat{y}^l - y^l) \right) \right).$$

$(\hat{y}^l - y^l)$ and $\delta^k(x, y)$ are bounded by the diameter $\alpha(x)$ (remember that in taking the moments we can substitute the pair $(\tilde{f}_x, \text{supp}(\tilde{f}_x))$ by $(f_x, \text{supp}(f_x))$ if we desire, since by *proposition (5.2.1)* the difference between the two distributions functions is small and because by *proposition (5.3.2)* we can replace the supports as well). Therefore,

$$\left\| \langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V} \right\|_{\tilde{\eta}} = \frac{1}{\text{vol}^2(\Sigma_x)} \left\| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \cdot \right. \right. \\ \left. \cdot \partial_j \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \delta^k(x, \hat{y}) \frac{\partial f}{\partial \hat{y}^l} (\hat{y}^l - y^l) \partial_k \right) \right\|_{\tilde{\eta}} \leq \\ \leq \frac{1}{\text{vol}^2(\Sigma_x)} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \frac{\partial f}{\partial \tilde{y}^l} \right) \right| \cdot \\ \cdot \left\| \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \delta^k(x, \hat{y}) (\hat{y}^l - y^l) \partial_k \right) \right\|_{\tilde{\eta}} + \\ + \frac{1}{\text{vol}^2(\Sigma_x)} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \frac{\partial f}{\partial \tilde{y}^l} \right) \right| \cdot \\ \cdot \left\| \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \partial_j \delta^k(x, \hat{y}) (\hat{y}^l - y^l) \partial_k \right) \right\|_{\tilde{\eta}} \leq \\ \leq \frac{1}{\text{vol}^2(\Sigma_x)} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \frac{\partial f}{\partial \tilde{y}^l} \right) \right| \cdot \\ \cdot \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \left\| \delta^k(x, \hat{y}) (\hat{y}^l - y^l) \partial_k \right\|_{\tilde{\eta}} \right) + \\ + \frac{1}{\text{vol}^2(\Sigma_x)} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \frac{\partial f}{\partial \tilde{y}^l} \right) \right| \cdot \\ \cdot \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \partial_j \left\| \delta^k(x, \hat{y}) (\hat{y}^l - y^l) \partial_k \right\|_{\tilde{\eta}} \right).$$

One can find a bound for each of these integrals. For instance, using the Hoelder inequality for integrals in an arbitrary space \mathbf{X} [13]

$$\left| \int_{\mathbf{X}} \lambda \phi \, d\mu \right| \leq \left(\int_{\mathbf{X}} |\lambda \, d\mu|^p \right)^{1/p} \left(\int_{\mathbf{X}} |\phi \, d\mu|^q \right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 \leq p, q \leq \infty.$$

We will use this inequality several times for the case $p = q = 2$, obtaining

$$\begin{aligned} & \left\| \left(\int_{\Sigma_x} d\text{vol}(x, \tilde{y}) (\tilde{y}^l - y^l) \delta^k(x, y) \partial_k \right) \right\|_{\bar{\eta}} \leq \\ & \leq \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) |(\hat{y}^l - y^l)| \left\| \delta^k(x, \hat{y}) \partial_k \right\|_{\bar{\eta}} \right) \leq \\ & \leq \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) |(\hat{y}^l - y^l)|^2 \right)^{\frac{1}{2}} \cdot \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \left\| \delta^k(x, \hat{y}) \partial_k \right\|_{\bar{\eta}}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Note that the index l is contracted with a factor $\frac{\partial f}{\partial \tilde{y}^l}$. Therefore $y^l \frac{\partial f}{\partial \tilde{y}^l}$ is Lorentz invariant and it can be computed in any inertial system, in particular in the laboratory frame. If we do this computation on this frame, we can continue with the above bound in the following way:

$$\begin{aligned} & \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) |(\hat{y}^l - y^l)|^2 \right)^{\frac{1}{2}} \cdot \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \left\| \delta^k(x, \hat{y}) \partial_k \right\|_{\bar{\eta}}^2 \right)^{\frac{1}{2}}. \\ & \leq \text{vol}_E^{\frac{1}{2}}(\Sigma_x) \alpha \cdot \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \left\| \delta^k(x, \hat{y}) \partial_k \right\|_{\bar{\eta}}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

In order to bound the second factor we use the following argument (that we used already in *section 4.6*),

$$\|\delta(x, y)\|_{\bar{\eta}} \leq \| \langle \hat{y} \rangle(x) - y \|_{\bar{\eta}} \leq \|\epsilon + \hat{y} - y\|_{\bar{\eta}} \leq \|\epsilon\|_{\bar{\eta}} + \|\hat{y} - y\|_{\bar{\eta}} \leq \frac{1}{2} \alpha + \alpha = \frac{3}{2} \alpha.$$

\hat{y} is in the support of the distribution f . Therefore, a bound on the integral is

$$\left\| \left(\int_{\Sigma_x} d\text{vol}(x, \tilde{y}) (\tilde{y}^l - y^l) \delta^k(x, \tilde{y}) \partial_k \right) \right\|_{\bar{\eta}} \leq \frac{3}{2} \cdot \text{vol}_E(\Sigma_x) \cdot \alpha^2.$$

Similarly one obtains the following bound:

$$\begin{aligned}
\left\| \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \partial_j(\delta^k(x, y)) (\hat{y}^l - y^l) \partial_k \right) \right\|_{\bar{\eta}} &\leq \left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) |(\hat{y}^l - y^l)|^2 \right)^{\frac{1}{2}} \\
&\cdot \left(\int_{\Sigma_x} d\text{vol}(x, y) \left\| \partial_j(\delta^k(x, y)) \partial_k \right\|_{\bar{\eta}}^2 \right)^{\frac{1}{2}} \leq \\
&\leq \text{vol}_E^{\frac{1}{2}}(\Sigma_x) \alpha \cdot \left(\int_{\Sigma_x} d\text{vol}(x, y) \left\| \partial_j(\delta^k(x, y)) \partial_k \right\|_{\bar{\eta}}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Then because of the definition of the corresponding Sobolev norm $\|\cdot\|_{0,2}$:

$$\begin{aligned}
\left(\int_{\Sigma_x} d\text{vol}(x, \hat{y}) \left\| \partial_j(\delta^k(x, \hat{y})) (\hat{y}^l - y^l) \partial_k \right\|_{\bar{\eta}} \right) &\leq \text{vol}_E^{\frac{1}{2}}(\text{supp}(\tilde{f}_x)) \alpha \cdot \|\partial_j \delta_x^k\|_{0,2} = \\
&= \text{vol}_E^{\frac{1}{2}}(\Sigma_x) \alpha \cdot \sum_k \|\delta_x^k \partial_j \log(\delta_x^k)\|_{0,2} \leq \text{vol}_E^{\frac{1}{2}}(\Sigma_x) \alpha^2 \cdot \sum_k \|\partial_j \log(\delta_x^k)\|_{0,2}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \frac{\partial f}{\partial \tilde{y}^l} \right) \right| &\leq \sum_{j=0}^{n-1} \left(\int_{\Sigma_x} d\text{vol}(x, y) |y^j \tilde{f}(x, y)|^2 \right)^{\frac{1}{2}} \\
&\cdot \left(\int_{\Sigma_x} d\text{vol}(x, y) \left| \partial_j \frac{\partial \tilde{f}(x, y)}{\partial y^k} \right|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

The second factor is equal to the Sobolev norm $\|\partial_j \tilde{f}_x\|_{1,1}$. The first factor is bounded in the following way:

$$\begin{aligned}
\left(\int_{\Sigma_x} d\text{vol}(x, y) |y^j \tilde{f}(x, y)|^2 \right)^{\frac{1}{2}} &\leq \left(\int_{\Sigma_x} d\text{vol}(x, y) |y^j|^2 \tilde{f}(x, y) \right)^{\frac{1}{2}} = \\
&= \text{vol}(\Sigma_x) \cdot (< |y^j|^2 >)^{\frac{1}{2}}.
\end{aligned}$$

Therefore, we get the bound:

$$\left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \frac{\partial f}{\partial \tilde{y}^l} \right) \right| \leq \text{vol}(\Sigma_x) \left(\sum_{j=0}^{n-1} (< |y^j|^2 >)^{\frac{1}{2}} \right) \|\partial_j \tilde{f}_x\|_{1,1}.$$

In a local frame where the vector field $U = (U^0, \vec{0})$, this contraction can be re-written

as

$$\begin{aligned} (< (y^0)^2 >)^\frac{1}{2} \cdot \|\partial_0 \tilde{f}_x\|_{1,1} &= \|(< |y^0|^2 >)^\frac{1}{2} \cdot \partial_0 \tilde{f}_x\|_{1,1} = \|(< (y^0)^2 >)^\frac{1}{2} \cdot \partial_0 \tilde{f}_x\|_{1,1} = \\ &= \|(< (y^0)^2 \cdot (\partial_0 \tilde{f}_x)^2 >)^\frac{1}{2}\|_{1,1} = \|(< (y^j \cdot \partial_j \tilde{f}_x)^2 >)^\frac{1}{2}\|_{1,1}. \end{aligned}$$

The last expression is covariant. Using normal coordinates associated with the affine connection $< {}^L \nabla >$ we obtain $\|(< (y^j \cdot \partial_j \tilde{f}_x)^2 >)^\frac{1}{2}\|_{1,1} = 0$.

Finally, we can bound the following integral

$$\begin{aligned} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \frac{\partial f}{\partial y^l} \right) \right| &\leq \left(\int_{\Sigma_x} d\text{vol}(x, y) |y^j \tilde{f}(x, y) \frac{\partial f}{\partial y^l}| \right) \leq \\ &\leq \sum_{j=0}^{n-1} \left(\int_{\Sigma_x} d\text{vol}(x, y) |y^j \tilde{f}(x, y)|^2 \right)^\frac{1}{2} \cdot \left(\int_{\Sigma_x} d\text{vol}(x, y) \left| \frac{\partial \tilde{f}(x, y)}{\partial y^k} \right|^\frac{1}{2} \right)^\frac{1}{2}. \end{aligned}$$

As in the previous integral, we get

$$\left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \frac{\partial f}{\partial y^l} \right) \right| \leq \text{vol}(\Sigma_x) \cdot (< |y^j|^2 >)^\frac{1}{2} \cdot \|\tilde{f}_x\|_{1,1}.$$

Using these bounds, we obtain the following relation:

$$\begin{aligned} \| < {}^L D >_{\tilde{V}} \tilde{V}(x) \|_{\bar{\eta}} &\leq \frac{1}{\text{vol}^2(\Sigma_x)}. \\ \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \partial_j \frac{\partial f}{\partial y^l} \right) \right| \cdot \frac{3}{2} \text{vol}_E(\text{supp}(\tilde{f}_x)) \alpha^2 &+ \\ + \frac{1}{\text{vol}^2(\Sigma_x)} \sum_{j=0}^{n-1} \left| \left(\int_{\Sigma_x} d\text{vol}(x, y) y^j \tilde{f}(x, y) \frac{\partial f}{\partial y^l} \right) \right| &\cdot \text{vol}_E^\frac{1}{2}(\Sigma_x) \cdot \alpha^2 \cdot \left(\sum_k \|\partial_j \log(\delta_x^k)\|_{0,2} \right) \leq \\ &\leq \frac{1}{\text{vol}^2(\Sigma_x)} \cdot \text{vol}(\Sigma_x) \cdot \\ \left(\sum_{j=0}^{n-1} (< |y^j|^2 >)^\frac{1}{2} \cdot \|\tilde{f}_x\|_{1,1} \cdot \text{vol}_E^\frac{1}{2}(\Sigma_x) \cdot \alpha^2 \cdot \left(\sum_k \|\partial_j \log(\delta_x^k)\|_{0,2} \right) \right) &= \end{aligned}$$

$$= \frac{vol_E^{\frac{1}{2}}(\Sigma_x)}{vol(\Sigma_x)} \cdot \left(\sum_{j=0}^{n-1} \langle |y^j|^2 \rangle^{\frac{1}{2}} \left(\sum_k \|\partial_j \log(\delta_x^k)\|_{0,2} \right) \right) \cdot \|\tilde{f}_x\|_{1,1} \cdot \alpha^2.$$

In a local frame where the vector field $U = \langle Y \rangle$ has components $(U^0, \vec{0})$, the following relation holds:

$$\begin{aligned} \sum_{j=0}^{n-1} \langle |y^j|^2 \rangle^{\frac{1}{2}} \left(\sum_k \|\partial_j \log(\delta_x^k)\|_{0,2} \right) &= \langle |y^0|^2 \rangle^{\frac{1}{2}} \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) = \\ &= \sum_k \langle (y^0)^2 \rangle \|\partial_0 \log(\delta_x^k)\|_{0,2}. \end{aligned}$$

Note that the normal coordinate system (that we are using) coincides with the adapted coordinate system, associated with the vector field $U = (U^0, \vec{0})$. In this coordinate system, there is a bound $\langle y^0 \rangle \leq 1 + \tilde{\alpha}$, where $\tilde{\alpha}$ is the diameter measured in the co-moving frame. It is of order 1 or smaller than 1. Therefore,

$$\sum_{j=0}^{n-1} \langle |y^j|^2 \rangle^{\frac{1}{2}} \left(\sum_k \|\partial_j \log(\delta_x^k)\|_{0,2} \right) \leq \langle |y^0|^2 \rangle^{\frac{1}{2}} \cdot \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) \cdot (1 + \tilde{\alpha}).$$

Then we have the following result:

$$\| \langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V}(x) \|_{\tilde{\eta}} \leq \frac{vol_E^{\frac{1}{2}}(\Sigma_x)}{vol(\Sigma_x)} \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) \cdot \|\tilde{f}_x\|_{1,1} \cdot \alpha^2 + \mathcal{O}(\alpha^3).$$

□

Corollary 5.3.4 *For compact domains $\mathbf{K} \subset \mathbf{M}$ and under the same hypotheses as in theorem 4.3, the following relation holds:*

$$\| \langle {}^L \nabla \rangle_{\tilde{V}} \tilde{V}(x) \|_{\tilde{\eta}} \leq n \cdot \tilde{C}(\mathbf{K}) \cdot \alpha^2 + \mathcal{O}(\alpha^3),$$

for some constant $\tilde{C}(\mathbf{K})$.

Proof: Take the constant $\tilde{C}(\mathbf{K})$ to be

$$\tilde{C}(\mathbf{K}) = \max_{x \in \mathbf{K}} \left\{ \frac{vol_E^{\frac{1}{2}}(\Sigma_x)}{vol(\Sigma_x)} \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) \cdot \|\tilde{f}_x\|_{1,1} \right\}.$$

□

These expressions are asymptotic formulas if $1 \gg \alpha$.

Remarks

1. In the preceding results the ultra-relativistic limit ($E \gg 1$) was not essential. However, the series in power of the energy has asymptotic meaning if $E \gg 1$.
2. There are several notions of normal coordinates, since we have several affine connections: ${}^\eta\nabla$, ${}^{\bar{\eta}}\nabla$ and $\langle {}^L\nabla \rangle$. However, we have only used the normal coordinates associated with $\langle {}^L\nabla \rangle$.

5.3.2 Bound on the auto-parallel condition of the unitary mean vector field of the averaged Vlasov model

Let us consider the normalized mean velocity vector field:

$$\tilde{u} = \frac{\tilde{V}}{\eta(\tilde{V}, \tilde{V})^{1/2}}.$$

Since $\langle {}^L\nabla \rangle$ does not preserve the Minkowski metric η , the covariant derivative of \tilde{u} in the direction of \tilde{u} using the Lorentz connection LD is

$$\langle {}^L\nabla \rangle_{\tilde{u}} \tilde{u} = \frac{1}{\eta(\tilde{V}, \tilde{V})} \langle {}^L\nabla \rangle_{\tilde{V}} \tilde{V} + \frac{1}{2} \left(\tilde{V} \cdot (\log(\eta(\tilde{V}, \tilde{V}))) \right) \tilde{V}. \quad (5.3.6)$$

The first term is bounded by *theorem 5.3.3*, since $\eta(\tilde{V}, \tilde{V}) > 1$. The total derivative of $\eta(\tilde{V}, \tilde{V})$ along a trajectory of \tilde{V} is

$$\begin{aligned} \mathcal{L}_{\tilde{V}}(\eta(\tilde{V}, \tilde{V})) &= \tilde{V} \cdot (\eta(\tilde{V}, \tilde{V})) = \\ &= 2\eta(\langle {}^L\nabla \rangle_{\tilde{V}} \tilde{V}, \tilde{V}) + (\langle {}^L\nabla \rangle_{\tilde{V}} \eta)(\tilde{V}, \tilde{V}). \end{aligned}$$

We have proved that the first term is of order α^2 . Using normal coordinates for $\langle {}^L\nabla \rangle$, one can compute the second term:

$$(\langle {}^L\nabla \rangle_{\tilde{V}} \eta)(\tilde{V}, \tilde{V}) = \eta(\tilde{V}, \tilde{V}) \mathbf{F}_{jm} \langle \delta^m(x, y) \delta^s(x, y) \delta^l(x, y) \rangle \tilde{V}^j \tilde{V}_s \tilde{V}_l.$$

We can estimate these contributions

Proposition 5.3.5 *Under the same assumptions as in theorem 5.3.3, the following relation holds:*

$$\langle {}^L \nabla \rangle_{\tilde{u}} \tilde{u} \leq \frac{vol_E^{\frac{1}{2}}(\Sigma_x)}{vol(\Sigma_x)} \left(\sum_k \|\partial_0 \log(\delta_x^k)\|_{0,2} \right) \cdot \|\tilde{f}_x\|_{1,1} \cdot \alpha^2 + \mathcal{O}(\alpha^3). \quad (5.3.7)$$

Proof: The first term of the right hand of the equation 5.3.6 is bounded by *theorem 5.3.3*. The second term is bounded using Hoelder's inequality for integrals [42, pg 62]

$$\left| \int_{\mathbf{X}} dvol(z) f_1(z) \cdots f_m(z) \right| \leq \prod_{k=1}^m \left(\int_{\mathbf{X}} dvol(z) |f_k(z)|^{p_k} \right)^{\frac{1}{p_k}}, \quad \sum_k p_k = 1, \quad 1 \leq p_k \leq \infty.$$

In particular one can apply this inequality to the third order moment

$$\langle \delta^m(x, y) \delta^s(x, y) \delta^l(x, y) \rangle:$$

$$\begin{aligned} \left| \langle \delta^m(x, y) \delta^s(x, y) \delta^l(x, y) \rangle \right| &= \frac{1}{vol(\Sigma_x)}. \\ &\cdot \left| \int_{\Sigma_x} dvol(x, y) f(x, y) \delta^m(x, y) \delta^s(x, y) \delta^l(x, y) \right| \\ &\leq \frac{1}{vol(\Sigma_x)} \left(\int_{\Sigma_x} dvol(x, y) |\tilde{f}(x, y) \delta^m(x, y)|^3 \right)^{\frac{1}{3}}. \\ &\quad \cdot \left(\int_{\Sigma_x} dvol(x, y) |\tilde{f}(x, y) \delta^s(x, y)|^3 \right)^{\frac{1}{3}}. \\ &\quad \cdot \left(\int_{\Sigma_x} dvol(x, y) |\tilde{f}(x, y) \delta^l(x, y)|^3 \right)^{\frac{1}{3}}. \end{aligned}$$

The distribution function is positive on $supp(\tilde{f})$. Also one can choose a distribution function such that $\tilde{f}_x \leq 1$. By proposition (5.3.1), one can substitute in the integrations $\tilde{f}_x \rightarrow f_x$, which implies that

$$\left| \langle \delta^m(x, y) \delta^s(x, y) \delta^l(x, y) \rangle \right| = O(\alpha^3).$$

Since the norm $\bar{\eta}(\langle y \rangle, \langle y \rangle) \geq 1$, one gets a third degree monomial term in α for the covariant derivative $\langle {}^L D \rangle_{\tilde{u}} \tilde{u}$. \square

Corollary 5.3.6 *Under the same assumptions as in theorem 5.3.3 in a compact domain $\mathbf{K} \subset \mathbf{M}$, one obtains*

$$\langle {}^L \nabla \rangle_{\tilde{u}} \tilde{u} \leq \cdot \tilde{C}(\mathbf{K}) \cdot \alpha^2 + \mathcal{O}(\alpha^3), \quad (5.3.8)$$

for a convenient constant $\tilde{C}(\mathbf{K})$.

5.3.3 Bound on the auto-parallel condition of the mean velocity field of the Vlasov model

Let us consider a local *Lorentz congruence*, which is a set of auto-parallel curves of the Lorentz connection ${}^L \nabla$, for a set of initial conditions at each (t_0, \vec{x}) , $\vec{x} \in \mathbf{M}_{t_0}$ where $\mathbf{M}_{t_0} \hookrightarrow \mathbf{M}$ is a 3-dimensional spatial sub-manifold. One can consider in a similar way the congruence associated with the averaged Lorentz connection for the same initial conditions. Note that, while the Lorentz connection preserves the Lorentz norm $\eta(\dot{x}, \dot{x})$ of the tangent vectors of the geodesics, this is not the case for the averaged Lorentz connection.

Theorem 5.3.7 *Let \mathbf{F} be a closed 2-form and ${}^L \nabla$ the associated non-linear Lorentz connection. Under the same assumptions as in theorem 5.3.3 the solutions of the Lorentz force equation $\eta \nabla_{\dot{x}} \dot{x} = (\iota_{\dot{x}} \mathbf{F})^\sharp$ can be approximated by the integral curves of the normalized mean velocity vector field $u = \frac{V(x)}{\sqrt{\eta(V(x), V(x))}}$ of the distribution function $f(x, y)$, where $f(x, y)$ is a solution of the associated Vlasov equation ${}^L \chi f = 0$. The difference is controlled by polynomial functions at least of order 2 in α ,*

$$\| {}^L \nabla_u u \|(x) \leq a_2(x) \alpha^2 + \mathcal{O}(\alpha^3) \quad (5.3.9)$$

where the function $a_2(x)$ is a bounded function of x .

Proof: We repeat an argument that we have used before. By *proposition 5.3.1*, both distribution functions f and \tilde{f} , solutions of ${}^L \chi f = 0$ and $\langle {}^L \chi \rangle \tilde{f}$, are such

that

$$|f(t, x(t), \dot{x}(t)) - \tilde{f}(t, x(t), \dot{x}(t))| \leq (\tilde{C}(x) \|\mathbf{F}\|_{\tilde{\eta}} C_2^2(x) (1 + B_2(x)\alpha)) \alpha^2 E^{-2} t^2 + \\ + (\tilde{K}(x) \|\mathbf{F}\|_{\tilde{\eta}}(x) K_2^2(1 + D_2(x)\alpha)) \alpha^2 E^{-1} t.$$

Therefore, the corresponding mean velocity fields are nearby as well, because of the linearity of the averaging operation and because of the above relation. Then their corresponding integral curves and the associated local congruences are also similar. By *corollary 5.3.6*, for narrow distributions, the normalized mean field \tilde{u} associated with \tilde{f} is such that

$$\| \langle {}^L\nabla \rangle_{\tilde{u}} \tilde{u}(x) \|_{\tilde{\eta}} \leq \tilde{a}_2 \alpha^2 + \mathcal{O}(\alpha^3).$$

for some function $\tilde{a}_2(x)$. Remember that one can interpolate smoothly between the connections ${}^L\nabla$ and $\langle {}^L\nabla \rangle$. Therefore, locally, one can interpolate smoothly between their integral curves. Also, because of the smoothness of the solutions of the geodesic equations with respect to the parameter of interpolation, there is a function a_2 in a small open neighborhood of \mathbf{M} such that

$${}^L\nabla_u u \leq a_2(x) \alpha^2 + \mathcal{O}(\alpha^3).$$

□

5.4 Discussion

Theorem 5.3.7 shows when the charged cold fluid model is a good approximation to the Vlasov equation in the description of the dynamics of a collection of particles interacting with an external electromagnetic field, in the ultra-relativistic regime. It is interesting that we have obtained this result without using additional hypotheses on the higher moments of the distribution function, except that the distribution is narrow and smooth enough for our calculations (we need some smoothness conditions

in order to use Taylor expansion for the function $f(x, y)$ in the velocity coordinates. Indeed, it seems that one only requires weak differentiability in y).

One of the important hypothesis on which the calculation is physically relevant is the requirement that the diameter of the distribution α must be small in the laboratory frame. Also, the ultra-relativistic regime $E \gg 1$ is useful in order to have good estimates and holds in current particle accelerators.

There are some technical issues that we would like to mention briefly:

1. We have assumed that the distribution functions \tilde{f} and f are at least \mathcal{C}^1 in x . However, let us consider the Dirac delta distribution with support invariant by the flow of the Lorentz force,

$$f(x, y) = \Psi(x) \delta(y - V(x)). \quad (5.4.1)$$

Since the width of the distribution is zero, $\alpha = 0$. One can use this distribution as a solution of the Vlasov equation in 2-dimensional space-times, for a proper value of the function Ψ . This example and the fact that the bounds found in *section 5.3* are formulated using Sobolev norms suggest the possibility of generalize the results to bigger function spaces. The results to use in this case are Sobolev embedding theorems [41, 57, 58]. However, we are not investigating this question in this thesis.

2. The same method can be applied to other fluid equations. Depending on the specific bounds and parameters, one can decide which model is better in each particular situation.

5.4.1 On the validity of the truncation schemes in fluid models

Given a kinetic model, usually the Maxwell-Vlasov system of differential equations, one defines a fluid model in terms of the low moments (typically, first, second, third and fourth moments) of the distribution function f [8, 37-39]. All the higher moments are set equal to zero. Usually, the typical reasoning

is that with the low moments one can write down models that are consistent and explain a reasonable number of phenomena in plasma physics.

We can argue that the reason why these models work is that in some situations the underlying Vlasov model can be substituted by the averaged Vlasov model. Then the Vlasov model depends only on the first, second and third moments of the distribution function f . Therefore, as soon as the hypothesis of a given fluid model are compatible with the hypothesis of the approximation *Maxwell-Vlasov model* \longrightarrow *averaged Maxwell-Vlasov model*, fluid models whose dynamical fields can be written in terms of the first, second and third moment, are *equivalent* to the underlying averaged Vlasov model. The equivalence must be understood in an approximated way, since there is an approximation in this argument.

The variables that one considers in fluid models are the mean velocity field (2.1.6), the covariant kinetic energy-momentum tensor (2.1.7) and the covariant energy-momentum flux tensor (2.1.8). Therefore, one can propose the following

Definition 5.4.1 *Two kinetic models are equivalent if their corresponding mean velocity field, covariant kinetic energy-momentum tensor and covariant energy-momentum flux tensor are the same.*

Definition 5.4.2 *Two fluid models are the same if their corresponding mean velocity field, covariant kinetic energy-momentum tensor and covariant energy-momentum flux tensor are the same.*

We propose the following conjecture in the form of a *theorem*

Theorem 5.4.3 *If two kinetic models are equivalent, the corresponding fluid models are the same. If two fluid models are the same, the underlying kinetic models are equivalent up to the order of approximation of the kinetic model by the averaged kinetic model.*

The first implication is trivial. The second implication is true for the Vlasov model, as we have proved in this *chapter*.

Therefore, when one works with a kinetic model, there is an underlying equivalence class of fluid models. We can call this an *universal class*. The elements of an universal class are, by construction, fluid models of ultra-relativistic narrow distributions. Then it is a useful idea to consider for each class the simplest model possible. In practice, the simplest model will dismiss higher order moments.

Chapter 6

The Jacobi equation of the averaged Lorentz connection and applications in beam dynamics

6.1 Introduction

6.1.1 The Jacobi equation of an affine connection on \mathbf{M}

Let \mathbf{M} be an n -dimensional manifold. Given an affine connection ∇ on the tangent bundle $\pi : \mathbf{TM} \longrightarrow \mathbf{M}$, the auto-parallel curves of ∇ are the solutions $c : \mathbf{I} \longrightarrow \mathbf{M}$ of the system of differential equations

$$\nabla_T T = 0, \quad T = \frac{dc}{dt},$$

where t is an affine parameter of ∇ . The curvature tensor R of the affine connection ∇ is the tensor field defined by the expression

$$R : \Gamma\mathbf{TM} \times \Gamma\mathbf{TM} \times \Gamma\mathbf{TM} \longrightarrow \Gamma\mathbf{TM}$$

$$(X, Y, Z) \mapsto R(X, Y, Z) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z, \quad \forall X, Y, Z \in \Gamma \mathbf{TM}. \quad (6.1.1)$$

The curvature tensor has an associated family of curvature endomorphisms $\{R_x(X, Y), \quad X, Y \in \Gamma \mathbf{TM}, x \in \mathbf{M}\}$ defined by

$$R(X, Y) : \Gamma \mathbf{TM} \longrightarrow \Gamma \mathbf{TM}$$

$$Z \mapsto R_x(X, Y)Z = R(X, Y, Z)(x), \quad Z \in \Gamma \mathbf{TM}, \forall x \in \mathbf{M}.$$

A vector field J along the parameterized geodesic $c : \mathbf{I} \longrightarrow \mathbf{M}$ of an affine, torsion-free connection is a Jacobi field if it satisfies the Jacobi equation

$$\nabla_X \nabla_X J - R(X, J)X = 0, \quad X = \frac{dc}{dt}. \quad (6.1.2)$$

This equation can be re-written using a local frame $\{e_i, i = 0, 1, 2, \dots, n-1\}$. It corresponds to the system of the second order differential equations

$$\frac{D^2 J^i}{dt^2} - R^i{}_{jkm}(c(t)) J^k \frac{dX^j}{dt} \frac{dX^m}{dt} = 0, \quad (6.1.3)$$

where $\frac{DJ}{dt}$ is the covariant derivative along the reference geodesic and the curvature tensor is given by the expression

$$R^i{}_{jkm} = \partial_m \Gamma^i{}_{jk} - \partial_k \Gamma^i{}_{jm} + (\Gamma^r{}_{jk} \Gamma^i{}_{rm} - \Gamma^r{}_{jm} \Gamma^i{}_{rk}).$$

For the next fundamental results one can consult [35, *section 10.1*; 54, § 14]. It is well known that for an affine connection there are $2n$ linear independent Jacobi fields along any given central geodesic c . This fact is a consequence of the existence and uniqueness of the solutions to second order differential equations and their smoothness properties on the initial values. A Jacobi field is completely determined by the value of $J(0)$ and $\frac{DJ}{dt}(0)$.

A smooth map

$$C : (-\lambda, \lambda) \times \mathbf{I} \longrightarrow \mathbf{M} \quad t \longrightarrow c_s(t) \subset \mathbf{M}$$

is a geodesic variation of $c(t)$ if each $c_s(t) : (-\lambda, \lambda) \longrightarrow \mathbf{M}$ is an affine parameterized geodesic curve for all $s \in \mathbf{I}$. The variation vector field of a geodesic variation is the vector field along the curve c defined by $dC(\frac{\partial}{\partial s})$, with $dC : \mathbf{T}((-\lambda, \lambda) \times \mathbf{I}) \longrightarrow \mathbf{TM}$ the differential of the smooth function C . The vector field $[dC(\frac{\partial}{\partial s}), \frac{dX}{dt}(c(t))]$ along the curve $c(t)$ vanish. The variation vector field acts on an arbitrary smooth function as a derivation

$$dC(f) = \frac{\partial f(C_t(s))}{\partial x^k} \frac{\partial C^k}{\partial t}.$$

The geometric interpretation of a Jacobi field is obtained through the following [35, 54]:

Proposition 6.1.1 *Let \mathbf{M} be a manifold equipped with an affine torsion-free connection ∇ . Then each Jacobi field J is a variation vector field of a geodesic variation C . Conversely, any variation field of a geodesic variation C defines a Jacobi field.*

6.1.2 Jacobi equation for linear connections defined on the pull-back bundle $\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$

Let \mathbf{M} be a smooth n -dimensional manifold, $\mathbf{N} \hookrightarrow \mathbf{TM}$ a sub-bundle of the tangent bundle \mathbf{TM} and let us consider an Ehresmann connection defined on \mathbf{TN} .

In order to formulate a Jacobi equation for linear connections on $\pi^*\mathbf{TM} \longrightarrow \mathbf{N}$, we mimic the standard derivation of the Jacobi equation for affine connections [35]. The *bending term* is determined by the hh -curvature endomorphisms:

$$R(h(X), h(Y)) : \pi^*\mathbf{TM} \longrightarrow \pi^*\mathbf{TM}$$

$$\zeta \mapsto R(h(X), h(Y))\zeta = (\nabla_{h(X)}\nabla_{h(Y)} - \nabla_{h(Y)}\nabla_{h(X)} - \nabla_{h([X, Y])})\zeta,$$

$$\forall \zeta \in \Gamma\mathbf{TM},$$

where the horizontal lift h , $h(X^i \partial_i) = X^i(x) \frac{\delta}{\delta x^i}$ was introduced in *section 3.3*.

Similar to the case of affine connections, one obtains the Jacobi equation for linear connections on $\pi^* \mathbf{TM} \rightarrow \mathbf{N}$. A basic fact is that we require that the Jacobi vector field commutes with the vector field $X(t)$ along the curve $c(t)$, which means that

$$[J, X]|_{c(t)} = 0.$$

Then the torsion-free condition along the curve c is

$$\nabla_{h(J)} \pi^* X = \nabla_{h(X)} \pi^* J,$$

and the curvature endomorphism along the curve c is such that

$$R(h(X), h(J)) \zeta = (\nabla_{h(X)} \nabla_{h(J)} - \nabla_{h(J)} \nabla_{h(X)}) \zeta, \quad \zeta \in \Gamma \pi^* \mathbf{TM}.$$

We can compute the second covariant derivatives along the curve c ,

$$\nabla_{h(X)} \nabla_{h(X)} \pi^* J = \nabla_{h(X)} \nabla_{h(J)} \pi^* X = \nabla_{h(J)} \nabla_{h(X)} \pi^* X - R(h(X), h(J)) \pi^* X. \quad (6.1.4)$$

Let us assume that $X \in \Gamma \mathbf{TM}$ is a auto-parallel respect to ∇ , which means $\nabla_{h(X)} \pi^* X = 0$. One obtains the following second order differential equation for $J(t)$

$$\nabla_{h(X)} \nabla_{h(X)} \pi^* J - R(h(X), h(J)) \pi^* X = 0. \quad (6.1.5)$$

Definition 6.1.2 *A field $J(t)$ along the curve $c : \mathbf{I} \rightarrow \mathbf{M}$ satisfying equation (6.1.5) is a Jacobi field of ∇ . The corresponding vector field $\pi_2(J(t))$ along the curve is the associated vector field.*

We can identify the vector field $\pi_2(J(t))$ with $J(t)$.

Definition 6.1.3 *Let ∇ be a linear connection on the bundle $\pi^* \mathbf{TM}$. An auto-parallel variation of the auto-parallel curve $c : \mathbf{I} \rightarrow \mathbf{M}$ is a map $C : (-\lambda, \lambda) \times \mathbf{I} \rightarrow$*

\mathbf{M} such that for each value of the parameter s , the curve $c_s(t) := C(s, t)$ is an auto-parallel curve, $\nabla_{h(\dot{c}_s(t))} \pi^* \dot{c}_s(t) = 0$.

Proposition 6.1.4 *Let ∇ a linear connection on $\pi^*\mathbf{TM}$. The variation field along $c(t)$ of a variation $C(t, s)$ is a Jacobi field of ∇ .*

Proof: It is clear from the deduction of the Jacobi equation for ∇ . \square

Remark. For covariant derivatives such that they are zero along the vertical direction, the covariant derivatives does not depend on the particular lift of $X \in \mathbf{T}_x\mathbf{M}$ to $\mathbf{T}_u\mathbf{N}$. Therefore for those covariant derivatives the expression that one obtains in this case for the Jacobi equation is defined as before and is independent of the vertical component of the lift that we are using.

6.1.3 Physical Interpretation of the Jacobi equation

For an affine connection, the Jacobi field represents the deviation vector of a given trajectory from the reference geodesic. However, if the reference trajectory is observable, the assumption that the reference trajectory coincides with the central geodesic provides physical meaning to the Jacobi field. Let us consider a geodesic variation $C(s, t)$. Each of the geodesics $c_s(t) = C(s, t)$ corresponds to a possible trajectory for a charged point particle. Then the Jacobi field corresponds to the deviation variable from a particular trajectory follow by a particle to the reference trajectory.

The above property holds at least for affine connections. Therefore, let us fix a central geodesic $c(t)$ and consider the set of all geodesic variations of the central geodesic $c(t)$. Hence there is a relation between the set of geodesic variations (which is equivalent to the set of Jacobi fields along $c(t)$ by *proposition 6.1.1*) and the set of all the trajectories allowable by the dynamics and the topology of the space-time manifold \mathbf{M} .

If the Jacobi vector field is given by $J(t)$, one identifies the components of $J(t)$ with the relative coordinates of a given geodesics respect the central affine geodesic $c(t)$, $J^k(t) = u^k(t)$.

Assuming this interpretation, the Jacobi Equation (for both affine and non-affine connections) is a second order Riccati equation

$$\frac{d^2 u}{dt^2} + R(t)u = 0.$$

This type of equation appears when one considers small deviations from a solution of another differential equation. One example is Hill's equation in celestial mechanics [55].

However, in the case where the connection is affine, the form of the Riccati equation is the same, but the endomorphism $R(t)$ along the curve $c(t)$ is simpler, since the curvature endomorphism depends only on the point $c(t)$ and not on the derivative $\frac{dc(t)}{dt}$ as in the general case of a non-linear connection.

6.2 Jacobi equation of the averaged lorentz connection

In this *section* \mathbf{M} is a n -dimensional manifold. The averaged Lorentz connection $\langle {}^L\nabla \rangle$ is an affine connection on \mathbf{M} .

Using local coordinates, a Jacobi field along the central geodesic can be written as $J = \xi^j(s)\partial_j$. The reference trajectory will be $X(\tau)$, which will be assumed to be a geodesic of the averaged Lorentz connection. Then the Jacobi equation for an affine connection on the tangent bundle $\mathbf{TM} \rightarrow \mathbf{M}$ can be expressed as

$$\frac{d^2 \xi^i}{d\tau^2} + 2\Gamma^i{}_{jk}(X(\tau)) \frac{d\xi^j}{d\tau} \frac{dX^k}{d\tau} + \xi^l \partial_l \Gamma^i{}_{jk}(X) \frac{dX^j}{d\tau} \frac{dX^k}{d\tau} = 0.$$

This equation is called the geodesic deviation equation. The central geodesic is denoted by $X(\tau)$ and a neighborhood geodesic is given by $x(\tau) = \xi(\tau) + X(\tau)$. The parameter τ is the proper-time of the central geodesic measured with the metric η .

The averaged Lorentz connection $\langle {}^L\nabla \rangle$ is an affine connection on the tangent bundle \mathbf{TM} . Therefore, we can apply the standard Jacobi equation to the averaged Lorentz connection. Given an arbitrary semi-Randers space $(\mathbf{M}, \eta, [A])$, in a local natural coordinate system, the averaged Lorentz connection has the connection

coefficients

$$\begin{aligned} < {}^L\Gamma^i{}_{jk} > = \eta\Gamma^i{}_{jk} + \frac{1}{2}(\mathbf{F}^i{}_j < y^m > \eta_{mk} + \mathbf{F}^i{}_k < y^m > \eta_{mj}) + \\ & + \mathbf{F}^i{}_m (< y^m > \eta_{jk} - \eta_{js}\eta_{kl} < y^m y^s y^l >), \end{aligned}$$

with $\mathbf{F} = dA$, with A being a representative of $[A]$, $A \in [A]$. The tangent vector y are on the unit hyperboloid $\Sigma_{\mathbf{x}}$.

The Jacobi equation of the averaged Lorentz connection is

$$\begin{aligned} & \frac{d^2 \xi^i}{d\tau^2} + 2 \frac{d\xi^j}{d\tau} \frac{dX^k}{d\tau} \left(\frac{1}{2}(\mathbf{F}^i{}_j < y^m > \eta_{mk} + \mathbf{F}^i{}_k < y^m > \eta_{mj}) + \right. \\ & + \mathbf{F}^i{}_m (< y^m > \eta_{jk} - \eta_{js}\eta_{kl} < y^m y^s y^l >)) + 2\xi^l \partial_l \left(\frac{1}{2}(\mathbf{F}^i{}_j < y^m > \eta_{mk} + \mathbf{F}^i{}_k < y^m > \eta_{mj}) + \right. \\ & + \mathbf{F}^i{}_m (< y^m > \eta_{jk} - \eta_{js}\eta_{kl} < y^m y^s y^l >)) \frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + \\ & + (\eta\Gamma^i{}_{jk} + \xi^l \partial_l \eta\Gamma^i{}_{jk}) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{ds} \frac{d\xi^k}{d\tau} \right) = 0. \end{aligned} \quad (6.2.1)$$

From the form of the system of differential equations (6.2.1) we conclude that:

1. There is a term representing the inertial acceleration:

$$\mathcal{A}_I := (\eta\Gamma^i{}_{jk} + \xi^l \partial_l \eta\Gamma^i{}_{jk}) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right). \quad (6.2.2)$$

The inertial acceleration $\mathcal{A}_{\mathcal{I}}$ is universal, in the sense that it is independent of the particle mass.

2. We can not say that $\mathcal{A}_{\mathcal{I}}$ is independent of the electromagnetic field, since $\frac{dX^j}{d\tau}$ can depend implicitly on the electromagnetic field when defining the reference trajectory. The typical example is the reference orbit of a betatron [9, *chapter 3*].

These properties of the inertial acceleration $\mathcal{A}_{\mathcal{I}}$ will help us to write it without doing explicit calculations, using establish formulae for elementary cases.

6.2.1 The Jacobi equation of the Lorentz connection versus the Jacobi equation of the averaged Lorentz connection

We have shown that the deviation equation from a given reference trajectory defines a Jacobi equation. However the non-linearity of the Lorentz force equation creates difficulties in view of the applicability of the above interpretation:

1. The evaluation of the covariant derivatives with respect to the original Lorentz connection requires a *reference vector*. Due to the dependence of the connection coefficients $\Gamma^i_{jk}(x, y)$ on the direction y , there must be assigned a particular point in the tangent space $y_0 \in \mathbf{T}_x\mathbf{M}$ at which the connection coefficients should be evaluated. This implies a specification of the direction where the connection coefficients are evaluated. Attaching a physical significance to the choice of the vector y_0 implies the selection of a particular model, which requires additional justification. This difficulty is resolved using the averaged Lorentz connection, which is an affine connection and whose connection coefficients do not depend on the *reference vector* y_0 .
2. It was proved in [22] (although in the category of Finsler spaces and for connections which covariant derivative vanish along vertical directions) that the averaged curvature of the original connection is the curvature of the averaged connection. This result can be extended to arbitrary linear connections on $\pi^*\mathbf{TM}$ with vanishing covariant derivative in the vertical directions. Hence, we can apply this result to the Lorentz connection. In the corresponding averaged Jacobi equation appears the averaged curvature, which is the same as the curvature of the averaged connection. Therefore we can work with the Jacobi equation for the averaged connection, as an attempt to give an *averaged description* of the dynamics of a beam of particles where the bending term is the *averaged force*.

Motivated by the above reasons, in this chapter we replace the Jacobi equation of the Lorentz connection by the Jacobi equation of the averaged Lorentz connection as a description of beam dynamics of bunches of particles in accelerators. We consider

systems in the ultra-relativistic regime and we also assume that the distribution function f is narrow in velocity space. We show how the Jacobi equation for the averaged connection $\langle {}^L\nabla \rangle$ provides a geometrical formulation of the *transversal* (in the case of dipole and quadrupole fields) and *longitudinal beam dynamics* (when the external fields are linearizable in the relative coordinates). We then provide a method to introduce corrections to the averaged Lorentz dynamics caused by the composed nature of the bunch of particles. These corrections are expressed in terms of known or observable quantities.

6.3 Transversal beam dynamics from the Jacobi equation of the averaged connection

Let us assume that (\mathbf{M}, η) is Minkowski space with \mathbf{M} being 4-dimensional. There is a global coordinate system denoted by (τ, x^1, x^2, x^3) . τ is the proper time considered from a given initial point of the reference trajectory, the coordinate x^2 is given by the Euclidean length of the path of the reference trajectory measured from the initial position in the reference frame defined by the vector field $\frac{d}{dt}$, which corresponds to the laboratory frame, x^1 is the horizontal coordinate and x^3 the vertical coordinate respect to the central geodesic. The longitudinal direction at each instant τ is given by the vector $\frac{\partial}{\partial x^2}$. By definition (x^1, x^3) are the transverse coordinates, while x^2 is the longitudinal coordinate.

6.3.1 Relation between the transverse dynamics and the Jacobi equation

We choose the laboratory reference frame for our calculations. Under the transverse dynamics, the difference $\frac{dx^2}{dt} - \frac{dX^2}{dt}$ will be constant. The external electromagnetic fields are static magnetic fields in Minkowski space. In the subsequent calculations we will only consider the lower order terms in the degree $a + b + c$ of the monomials $\xi^a (\frac{d\xi}{d\tau})^b \epsilon^c$ appearing in the expressions, with $\xi = x(\tau) - X(\tau)$, $\epsilon = \langle y \rangle - \frac{dX}{d\tau}$.

Firstly, we linearize the equations with respect to the degree defined by the vector

fields along the central geodesic ξ and the powers of the difference ϵ appearing on each term. Recall that the *transverse component*, which are the terms in the geodesic equation proportional to the tensor $T^i_{jk} = \mathbf{F}^i_m (\langle y^m \rangle \eta_{jk} - \eta_{js} \eta_{kl} \langle y^m y^s y^l \rangle)$, are neglected systematically because they are of higher order in the degree $(a+b+c)$ than the *longitudinal component*. The longitudinal component is proportional to

$$L^i_{jk} = \frac{1}{2} (\mathbf{F}^i_j \langle y^m \rangle \eta_{mk} + \mathbf{F}^i_k \langle y^m \rangle \eta_{mj})$$

After linearizing, the system of differential equations are

$$\begin{aligned} & \frac{d^2 \xi^i}{d\tau^2} + 2 \frac{d\xi^j}{d\tau} \frac{dX^k}{d\tau} \left(\frac{1}{2} (\mathbf{F}^i_j \langle y^m \rangle \eta_{mk} + \mathbf{F}^i_k \langle y^m \rangle \eta_{mj}) \right) + \\ & + 2 \xi^l \partial_l \left(\frac{1}{2} (\mathbf{F}^i_j \langle y^m \rangle \eta_{mk} + \mathbf{F}^i_k \langle y^m \rangle \eta_{mj}) \right) \cdot \frac{dX^j}{d\tau} \frac{dX^k}{d\tau} \\ & + (\eta \Gamma^i_{jk} + \xi^l \partial_l \eta \Gamma^i_{jk}) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0, \quad i, j, k, m = 0, \dots, n-1. \end{aligned}$$

Since ϵ is small, we can replace $\langle y \rangle \longrightarrow \frac{dX}{ds}$, obtaining the following differential equation:

$$\begin{aligned} & \frac{d^2 \xi^i}{d\tau^2} + 2 \frac{d\xi^j}{d\tau} \frac{dX^k}{d\tau} \left(\frac{1}{2} (\mathbf{F}^i_j \frac{dX^m}{d\tau} \eta_{mk} + \mathbf{F}^i_k \frac{dX^m}{d\tau} \eta_{mj}) \right) + \\ & + 2 \xi^l \partial_l \left(\frac{1}{2} (\mathbf{F}^i_j \frac{dX^m}{d\tau} \eta_{mk} + \mathbf{F}^i_k \frac{dX^m}{d\tau} \eta_{mj}) \right) \cdot \frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + \\ & + (\eta \Gamma^i_{jk} + \xi^l \partial_l \eta \Gamma^i_{jk}) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0. \end{aligned}$$

Remark. This equation is the geodesic deviation equation of the Lorentz force equation. Therefore, the difference between the averaged Jacobi equation and the deviation equation associated with the Lorentz force is of higher order in the degree $(a+b+c)$. At leading order both equations coincide.

The condition of transversal dynamics is

$$\frac{d\xi^j}{d\tau} \frac{dX_j}{d\tau} \simeq \mathcal{O}^2,$$

where \mathcal{O}^2 indicates first order in the general degree $(a+b+c)$. This is a more general

condition than what is usually stated as the transverse dynamics in accelerator physics: the magnetic field is perpendicular to the velocity field $\frac{dX}{d\tau}$ of the particles in the beam [9].

Due to this condition, we suppress the respective term in the differential equations, getting the equations

$$\begin{aligned} \frac{d^2\xi^i}{d\tau^2} + 2\frac{d\xi^j}{d\tau}\frac{dX^k}{d\tau}\left(\frac{1}{2}\mathbf{F}^i{}_j\frac{dX^m}{d\tau}\eta_{mk}\right) + 2\xi^l\partial_l\left(\frac{1}{2}\mathbf{F}^i{}_j\frac{dX^m}{d\tau}\eta_{mk}\right)\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + \\ + (\eta\Gamma^i{}_{jk} + \xi^l\partial_l\eta\Gamma^i{}_{jk})\left(\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau}\right) = 0. \end{aligned}$$

Therefore, at first order, the differential equations are

$$\frac{d^2\xi^i}{d\tau^2} + \frac{d\xi^j}{d\tau}\mathbf{F}^i{}_j + \frac{dX^j}{d\tau}\xi^l\partial_l\mathbf{F}^i{}_j + \eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + \xi^l\partial_l\eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau} = 0.$$

In the transverse dynamics, one assumes by construction that $\frac{d\xi^j}{d\tau}\mathbf{F}^i{}_j = 0$, if there is not *dispersion*. Hence, the differential equations in this regime are

$$\frac{d^2\xi^i}{d\tau^2} + \frac{d\xi^j}{d\tau}\mathbf{F}^i{}_j + \eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + \xi^l\partial_l\eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\eta\Gamma^i{}_{jk}\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau} = 0. \quad (6.3.1)$$

As we mentioned before, the last term of this equation corresponds to the inertial term. In a similar way as in the circular motion, the inertial terms is already well known [9,10]; the components of the *inertial acceleration* will be assumed to be

$$\begin{aligned} (\eta\Gamma^0{}_{jk} + \xi^l\partial_l\eta\Gamma^0{}_{jk})\left(\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau}\right) &= 0, \\ (\eta\Gamma^3{}_{jk} + \xi^l\partial_l\eta\Gamma^3{}_{jk})\left(\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau}\right) &= 0, \\ (\eta\Gamma^2{}_{jk} + \xi^l\partial_l\eta\Gamma^2{}_{jk})\left(\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau}\right) &= 0, \\ (\eta\Gamma^1{}_{jk} + \xi^l\partial_l\eta\Gamma^1{}_{jk})\left(\frac{dX^j}{d\tau}\frac{dX^k}{d\tau} + 2\frac{dX^j}{d\tau}\frac{d\xi^k}{d\tau}\right) &= \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\xi + \rho} - \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} = \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} \left(-\frac{\xi^1}{\rho}\right). \end{aligned}$$

In the last expression we consider ξ to be small in relation to the curvature radius of the central geodesic ρ . Note that ρ is not necessarily constant. However, we are

assuming a planar trajectory. In the usual formalism, the last term corresponds to the *relative centripetal force* between two particles following close trajectories.

6.3.2 Examples of transverse linear dynamics

We study some examples of transverse dynamics using the *linearized version* of the averaged Jacobi equation.

1. Motion in a normal magnetic dipole

As we said before, the reference frame is the laboratory frame. The reference trajectory is a solution of the averaged Lorentz force equation. In this case the electromagnetic field is given by the expression

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_0 & 0 \\ 0 & -b_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

where b_0 is the dipole strength. Since the magnetic field is constant, $\xi^l \partial_l \mathbf{F}^i_j = 0$. Therefore, the equations of motion for the transverse degrees of freedom (ξ^1, ξ^3) are

$$\frac{d^2 \xi^1}{d\tau^2} + \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} \left(-\frac{\xi^1}{\rho}\right) = 0, \quad \frac{d^2 \xi^3}{d\tau^2} = 0.$$

Changing the parameter of the curve from $\tau \rightarrow x^1$, one has that:

$$\frac{d^2 \xi^1}{dl^2} - \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{dl^2} = 0.$$

These are the standard equations for a normal dipole.

2. Motion in a skew magnetic dipole

The electromagnetic field is given by the expression

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -b_0 & 0 \\ 0 & b_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In this case, the Jacobi equations are

$$\frac{d^2 \xi^1}{d\tau^2} - \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} \left(-\frac{\xi^1}{\rho}\right) = 0, \quad \frac{d^2 \xi^3}{d\tau^2} = 0.$$

Following the same procedure as before we end with the equations for the deviation equation in a skew magnetic field:

$$\frac{d^2 \xi^1}{dl^2} - \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{dl^2} = 0.$$

3. Motion in a normal quadrupole field combined with a dipole

In this case the electromagnetic field has the form

$$\mathbf{F}(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_0 - b_1 \xi^1 & 0 \\ 0 & -b_0 + b_1 \xi^1 & 0 & b_1 \xi^3 \\ 0 & 0 & -b_1 \xi^3 & 0 \end{pmatrix}$$

The Jacobi equation reduces to

$$\frac{d^2 \xi^1}{d\tau^2} - \frac{dX^j}{d\tau} \xi^l \partial_l \mathbf{F}^1_j + \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} \left(-\frac{\xi^1}{\rho}\right) = 0,$$

$$\frac{d^2 \xi^3}{d\tau^2} + \frac{dX^j}{d\tau} \xi^l \partial_l \mathbf{F}^3_j = 0.$$

Let us consider the respective contributions $\frac{d\xi^j}{d\tau} \xi^l \partial_l \mathbf{F}^1_j$ and $\frac{d\xi^j}{ds} \xi^l \partial_l \mathbf{F}^3_j$. Using

Euler's theorem on homogenous functions one gets the relations:

$$\frac{dX^j}{d\tau} \xi^l \partial_l \mathbf{F}^3_j = \frac{dX^j}{d\tau} \mathbf{F}^3_j|_{\xi=0}.$$

Then the differential equations are

$$\frac{d^2 \xi^1}{d\tau^2} - \xi^1 b_1 + \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{1}{\rho} \left(-\frac{\xi^1}{\rho}\right) = 0, \quad \frac{d^2 \xi^3}{d\tau^2} + \xi^3 b_1 = 0.$$

Using the Euclidean length as a parameter of the curve, one obtains

$$\frac{d^2 \xi^1}{dl^2} - \xi^1 \frac{\partial B^3}{\partial \xi^1} + \left(\frac{d\vec{X}}{dl}\right)^2 \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{dl^2} + \frac{d\xi^2}{dl} b_0^1 + \xi^3 \frac{\partial B^1}{\partial \xi^3} = 0.$$

In the second differential equation, the second term is zero, since $(\mathbf{b}_0^1, \mathbf{b}_0^2, \mathbf{b}_0^3) = (0, 0, \mathbf{b}_0^3)$. Then we obtain the following differential equations for the transverse motion:

$$\frac{d^2 \xi^1}{d\tau^2} - \xi^1 b_1 + \left(\frac{d\vec{X}}{d\tau}\right)^2 \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{d\tau^2} + \xi^3 b_1 = 0.$$

These are the equations of the linear transverse dynamics in quadrupoles combined with magnetic dipole fields using the proper time parameter τ . If we use the Euclidean length l , the equations are

$$\frac{d^2 \xi^1}{dl^2} - \xi^1 b_1 + \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{dl^2} + \xi^3 b_0 = 0, \quad (6.3.2)$$

which are the standard equations in transverse dynamics in accelerator physics [9, 10].

4. Motion in a normal dipole combined with a 45 degrees quadrupole

In this case the electromagnetic field is

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & b_0 + b_1 \xi^3 & 0 \\ 0 & -b_0 - b_1 \xi^3 & 0 & b_1 \xi^1 \\ 0 & 0 & -b_1 \xi^1 & 0 \end{pmatrix}$$

Following the same procedure as before, we get the Jacobi equations

$$\frac{d^2 \xi^1}{dl^2} + \xi^1 b_1 + \frac{\xi^1}{\rho^2} = 0, \quad \frac{d^2 \xi^3}{dl^2} - \xi^3 b_0 = 0. \quad (6.3.3)$$

The above examples show how the linear transverse dynamics can be obtained from the Jacobi equation of the averaged Lorentz connection.

It is not possible to use this formulation for higher multipole magnetic fields because the linear approximation breaks down. One possibility to incorporate higher modes is to consider the generalized Jacobi equation [59], which is a non-linear geodesic deviation equation.

We remark again that with the approximation $\langle y \rangle \longrightarrow \frac{dX}{d\tau}$ the deviation equation of the averaged connection and the Lorentz connection coincide. If one considers higher order effects, one can obtain differences between the Jacobi equations of the Lorentz connections and averaged connection.

6.4 Calculation of the averaged off-set effect between the reference trajectory and the central geodesic of the averaged Lorentz connection

In this section we calculate the averaged difference between the reference trajectory and the solutions of the averaged connection.

6.4.1 Calculation of the dispersion function in beam dynamics

We follow the formalism developed in [9] for the treatment of linear perturbations and dispersion. However we will maintain the proper time τ as the parameter of the curves, in contrast with the usual treatment, which uses the Euclidean length along the reference trajectory. In the following, primes indicate derivatives with respect to the proper time.

It follows from *section 6.3* that the transverse dynamics is determined by equations

of the form

$$u'' + K(\tau)u = 0 \quad (6.4.1)$$

The general solution is of the form

$$u(\tau) = C(\tau)u_0 + S(\tau)u'_0, \quad u'(\tau) = C'(\tau)u_0 + S'(\tau)u'_0,$$

with initial conditions

$$C(0) = 1, \quad C'(0) = 0; \quad S(0) = 0, \quad S'(0) = 1,$$

for arbitrary initial values u_0 and u'_0 . The functions $C(\tau)$ and $S(\tau)$ satisfy

$$C''(\tau) + K(\tau)S(\tau) = 0, \quad S''(\tau) + K(\tau)C(\tau) = 0.$$

However, small perturbations can change the dynamics. The perturbed equation has the form:

$$u''(\tau) + K(\tau)u(\tau) = p(\tau) \quad (6.4.2)$$

A particular solution for (6.4.2) is

$$P(\tau) = \int_0^\tau p(\tilde{\tau})G(\tau, \tilde{\tau})d\tilde{\tau}, \quad (6.4.3)$$

where $G(\tau, \tilde{\tau})$ is the Green function associated to the differential equation (6.4.1). One can prove that in the absence of dissipative forces (that is, which do not depend on the velocity of the particle), the Green function of the differential equation is given by the following combination:

$$G(\tau, \tilde{\tau}) = S(\tau)C(\tilde{\tau}) - C(\tau)S(\tilde{\tau}). \quad (6.4.4)$$

Therefore, the general solution for the equation (6.4.2) is

$$u(\tau) = aC(\tau) + bS(\tau) + P(\tau). \quad (6.4.5)$$

This solution breaks down if there are synchrotron radiation or other dissipative effects.

We will use standard notation of beam dynamics. If all the particles in a bunch do not have the same energy, one obtains for the transverse degrees of freedom the following differential equation [9, pg 109], [10]:

$$u'' + K(\tau)u = \frac{1}{\rho_0}(\tau)\Delta_u, \quad \Delta = \frac{\delta p}{p_0}, \quad \delta p = \sqrt{(\delta p_1)^2 + (\delta p_2)^2 + (\delta p_3)^2}.$$

We need to assign a value to δp . One natural value is the maximal value of $\{\|\xi(\vec{x})\|_{\vec{\eta}}, x \in \mathbf{M}\}$. Since \mathbf{M} is non-compact, we restrict to a compact domain $\mathbf{K} \subset \mathbf{M}$. This definition does not depend on the particular trajectory of each particle. The general solution for u is linear in the perturbation, therefore

$$u(\tau) = a_u C(\tau) + b_u S(\tau) + \Delta D(\tau) := a_u C(\tau) + b_u S(\tau) + O f f_u(\tau), \quad P(\tau) = \Delta D(\tau).$$

where a and b depend on the initial values.

6.4.2 Calculation of the off-set due to the deviation ϵ^j

In this subsection (\mathbf{M}, η) is the 4-dimensional Minkowski space-time. We have shown in the previous section that in first order of approximation with respect to the degree $(a + b)$ of the monomials $\xi^a \cdot (\epsilon^m)^b$ and its derivatives, when we take the approximation $\langle y \rangle \longrightarrow \frac{dX}{dt}$, the differential equation for the transverse motion is the Jacobi equation of the averaged connection. Therefore we can consider the terms on ϵ^k in the averaged Jacobi equation as a perturbation and apply the method of the Green function.

From the definition of the off-set function for the transverse degrees of freedom, we obtain

$$O f f_u^{1,3}(\tau) = u^{1,3}(\tau) - a_{1,3} C^{1,3}(\tau) - b_{1,3} S^{1,3}(\tau),$$

the super-index refers to the transverse components x^2 and x^3 in the laboratory

frame defined previously. Using the corresponding Green function we obtain

$$Of f_u^{1,3}(\tau) = \int_0^\tau p^{1,3}(\tau) G(\tau, \tilde{\tau}) d\tilde{\tau}.$$

The perturbation $p(\tau)$ is in this case defined by all the terms of the averaged Jacobi equation which are not contained in the linearized equation respect to the degree $(a+b+c)$. Therefore let us re-write the Jacobi equation of the averaged connection.

Using $\epsilon_k = \langle y_k \rangle - \frac{dX_k}{d\tilde{\tau}}$ we get

$$\begin{aligned} Of f_u^{1,3}(\tau) = & \int_0^\tau d\tilde{\tau} 2 \frac{d\xi^j}{d\tilde{\tau}}(\tilde{\tau}) \cdot \frac{dX^k}{d\tilde{\tau}} \left(\frac{1}{2} (\mathbf{F}^{1,3}_{\ j}(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \mathbf{F}^{1,3}_{\ k}(\tilde{\tau}) \epsilon_j(\tilde{\tau})) + \right. \\ & + \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} \left(\mathbf{F}^{1,3}_{\ m}(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) + \right. \\ & \left. \left. + \xi^l \partial_l (\mathbf{F}^{1,3}_{\ m}(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk})) \right) \right). \end{aligned} \quad (6.4.6)$$

This is an integro-differential equation for $Of f_u^{1,3}$ as we can show. In the integrand of equation (6.4.6) we can make the substitution

$$\frac{d\xi^{1,3}}{d\tilde{\tau}}(\tilde{\tau}) \longrightarrow (a_{1,3} C'^{1,3}(\tilde{\tau}) + b_{1,3} S'^{1,3}(\tilde{\tau}) + Of f_u(\tilde{\tau})').$$

We can also consider the derivatives in the longitudinal and temporal direction using this notation, with a convenient choice of the coefficients $a_{0,2}$ and $b_{0,2}$. Then we can write

$$\frac{d\xi^j}{d\tilde{\tau}}(\tilde{\tau}) \longrightarrow (a_j C'^j(\tilde{\tau}) + b_j S'^j(\tilde{\tau}) + Of f_u^j(\tilde{\tau})'), \quad j = 0, 1, 2, 3.$$

In this expression repeated indices are not summed! For the transverse degrees of freedom, the unperturbed solutions are the same as before [10],

$$u^{1,3}(\tau) = u_0 C(\tau) + u'_0 S(\tau) \quad C''(\tau) + K(\tau)C = 0, \quad S''(\tau) + K(\tau)S = 0.$$

For the longitudinal $j = 2$ and temporal $j = 0$ degrees of freedom, one gets the

following relations by comparison with the Jacobi equation,

$$\begin{aligned}\frac{d\xi^2}{d\tilde{\tau}}(\tilde{\tau}) &\longrightarrow (a_2 C'^2(\tilde{\tau}) + b_2 S'^2(\tilde{\tau}) + Of f_u^2(\tilde{\tau})'), \\ \frac{d\xi^0}{d\tilde{\tau}}(\tilde{\tau}) &\longrightarrow (a_0 C'^0(\tilde{\tau}) + b_0 S'^0(\tilde{\tau}) + Of f_u^0(\tilde{\tau})')\end{aligned}$$

Let us consider the regime where $Of f_u^0 = Of f_u^2 = 0, \forall u$. Then the off-set function is

$$\begin{aligned}Of f_u^{1,3}(\tau) &= \int_0^\tau d\tilde{\tau} \left(\sum_{j=0}^3 2(a_j C'^j(\tilde{\tau}) + b_j S'^j(\tilde{\tau}) + Of f_u^j(\tilde{\tau})')(\tilde{\tau}) \cdot \frac{dX^k}{d\tilde{\tau}} \left(\frac{1}{2} (\mathbf{F}^{1,3}_j(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \right. \right. \\ &+ \mathbf{F}^{1,3}_k(\tilde{\tau}) \epsilon_j(\tilde{\tau})) + \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) + \right. \\ &+ (a_l C^l(\tilde{\tau}) + b_l S^l(\tilde{\tau}) + Of f_u^l(\tilde{\tau})) \partial_l \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^l \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) \right) \Big) \Big). \end{aligned}$$

This is an integro-differential equation for $Of f_u^{1,3}$ that we formally can solve iteratively. In *the Born approximation* one puts $Of f_u^l(\tilde{\tau}) = 0$ in the integrand:

$$\begin{aligned}Of f_\xi^{1,3}(\tau) &= \int_0^\tau d\tilde{\tau} \left(2(a_j C'^j(\tilde{\tau}) + b_j S'^j(\tilde{\tau})) \cdot \frac{dX^k}{d\tilde{\tau}} \left(\frac{1}{2} (\mathbf{F}^{1,3}_j(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \mathbf{F}^{1,3}_k(\tilde{\tau}) \epsilon_j(\tilde{\tau})) + \right. \right. \\ &+ \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) + \right. \\ &+ (a_l C^l(\tilde{\tau}) + b_l S^l(\tilde{\tau})) \partial_l \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^l \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) \right) \Big) \Big). \end{aligned}$$

Therefore, we get the expression

$$\begin{aligned}Of f_\xi^{1,3}(\tau) &= \int_0^\tau d\tilde{\tau} \left(2 \frac{d\xi^j}{d\tilde{\tau}} \cdot \frac{dX^k}{d\tilde{\tau}} \left(\frac{1}{2} (\mathbf{F}^{1,3}_j(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \mathbf{F}^{1,3}_k(\tilde{\tau}) \epsilon_j(\tilde{\tau})) \right. \right. \\ &+ \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) + \right. \\ &+ \xi^l \partial_l \left(\mathbf{F}^{1,3}_m(\tilde{\tau}) (\langle y^m \rangle(\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^l \rangle(\tilde{\tau}) \eta_{ja} \eta_{lk}) \right) \Big) \Big). \end{aligned} \quad (6.4.7)$$

This expression depends on the particular solution u . A way to eliminate this

dependence is to take the following *average*

$$\begin{aligned}
\langle Of f_u^{1,3} \rangle (\tau) = & \int_0^\tau d\tilde{\tau} \left(2\epsilon^j(\tilde{\tau}) \cdot \frac{dX^k}{d\tilde{\tau}} \left(\frac{1}{2} (\mathbf{F}^{1,3}_{\ j}(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \mathbf{F}^{1,3}_{\ k}(\tilde{\tau}) \epsilon_j(\tilde{\tau})) \right. \right. \\
& + \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} \left(\mathbf{F}^{1,3}_{\ m}(\tilde{\tau}) (\langle y^m \rangle (\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle (\tilde{\tau}) \eta_{ja} \eta_{lk}) + \right. \\
& \left. \left. + \langle \xi^l \rangle (\tilde{\tau}) \partial_l (\mathbf{F}^{1,3}_{\ m}(\tilde{\tau}) (\langle y^m \rangle (\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^l \rangle (\tilde{\tau}) \eta_{ja} \eta_{lk})) \right) \right). \quad (6.4.8)
\end{aligned}$$

The averaged off-set is therefore an observable quantity. It is determined by:

1. The reference trajectory $X(\tau)$, which is a geodesic of the averaged connection and as we have discussed before, it is known theoretically.
2. The tangent velocity field $\frac{dX}{d\tau}$ along the reference trajectory. This is known theoretically.
3. The external electromagnetic field $\mathbf{F}^{1,3}_{\ m}(x)$,
4. The value of the vector field $\epsilon^k(\tau) = \langle y^k(\tau) \rangle - \frac{dX}{d\tau}$,
5. The first, second and third moments of the distribution function $f(x(\tau), p(\tau))$ along the reference trajectory.

Finally, in the case that the perturbation does not change significantly along the trajectory, we obtain that the term containing derivatives are neglected. Therefore,

$$\begin{aligned}
\langle Of f_u^{1,3} \rangle (\tau) = & \int_0^\tau d\tilde{\tau} \langle \left(2\epsilon^j(\tilde{\tau}) \frac{dX^k}{d\tilde{\tau}} \cdot (\mathbf{F}^{1,3}_{\ j}(\tilde{\tau}) \epsilon_k(\tilde{\tau}) + \mathbf{F}^{1,3}_{\ k}(\tilde{\tau}) \epsilon_j(\tilde{\tau})) + \right. \\
& \left. + \frac{dX^j}{d\tilde{\tau}} \frac{dX^k}{d\tilde{\tau}} (\mathbf{F}^{1,3}_{\ m}(\tilde{\tau}) (\langle y^m \rangle (\tilde{\tau}) \eta_{jk} - \langle y^m y^a y^k \rangle (\tilde{\tau}) \eta_{ja} \eta_{lk})) \right) \rangle. \quad (6.4.9)
\end{aligned}$$

In the case of a delta function distribution we have $\langle Of f_u^{1,3} \rangle (\tau) = 0$. This means that the averaged off-set effect is a collective effect.

6.5 Longitudinal beam dynamics and corrections from the Jacobi equation of the averaged connection

Let (\mathbf{M}, η) be the Minkowski space-time and consider an inertial coordinate system defined by the vector field $Z = \frac{\partial}{\partial t}$, that corresponds to the laboratory frame. The interaction of an ultra-relativistic bunch of particles with an external longitudinal electric field is described by the Faraday tensor

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & E_2(x) & 0 \\ 0 & 0 & 0 & 0 \\ -E_2(x) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

For narrow distributions one obtains the following condition,

$$\frac{dX^j}{d\tau} \frac{d\xi_j}{d\tau} = \mathcal{O}^1.$$

This relation can be seen as follows. For the linear dynamics $\xi = (\xi, 0, -\xi, 0)$ in the laboratory frame. Using the ultra-relativistic limit $\frac{dX^k}{ds} = (1 + E, 0, E, 0)$, with $E \gg 1$.

Then the averaged Jacobi equation for the limit $\epsilon^j \rightarrow 0$ in the ultra-relativistic regime are

$$\begin{aligned} & \frac{d^2 \xi^i}{d\tau^2} + 2 \frac{d\xi^j}{d\tau} \frac{dX^k}{d\tau} \left(\frac{1}{2} (\mathbf{F}^i{}_j \frac{dX^m}{d\tau} \eta_{mk} + \mathbf{F}^i{}_k \frac{dX^m}{d\tau} \eta_{mj}) \right) + 2 \xi^l \partial_l \left(\frac{1}{2} (\mathbf{F}^i{}_j \frac{dX^m}{d\tau} \eta_{mk} + \mathbf{F}^i{}_k \frac{dX^m}{d\tau} \eta_{mj}) \right) \\ & \cdot \frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + (\eta \Gamma^i{}_{jk} + \xi^l \partial_l \eta \Gamma^i{}_{jk}) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0. \end{aligned}$$

For the above longitudinal electric field, the equations of motion are

$$\begin{aligned} & \frac{d^2 \xi^0}{d\tau^2} + \frac{d\xi^2}{d\tau} \frac{dX^k}{d\tau} E_2 < y^m > \eta_{mk} + \frac{d\xi^k}{d\tau} \frac{dX^2}{d\tau} E_2 < y^m > \eta_{mk} + \\ & + \xi^l \partial_l \left(\frac{dX^k}{d\tau} \frac{dX^2}{d\tau} E_2 < y^m > \eta_{mk} + \frac{dX^2}{d\tau} \frac{dX^j}{d\tau} E_2 < y^m > \eta_{mj} \right) + \end{aligned}$$

$$\begin{aligned}
& + (\eta \Gamma_{jk}^0 + \xi^l \partial_l \eta \Gamma_{jk}^0) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0, \\
& \frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^0}{d\tau} \frac{dX^k}{d\tau} E_2 < y^m > \eta_{mk} - \frac{d\xi^k}{d\tau} \frac{dX^0}{d\tau} E_2 < y^m > \eta_{mk} - \\
& \xi^l \partial_l \left(\frac{dX^j}{d\tau} \frac{dX^0}{d\tau} E_2 < y^m > \eta_{mj} + \frac{dX^0}{d\tau} \frac{dX^j}{d\tau} E_2 < y^m > \eta_{mj} \right) + \\
& + (\eta \Gamma_{jk}^2 + \xi^l \partial_l \eta \Gamma_{jk}^2) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0, \\
& \frac{d^2 X^1}{d\tau^2} + (\eta \Gamma_{jk}^1 + \xi^l \partial_l \eta \Gamma_{jk}^1) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0, \\
& \frac{d^2 X^2}{d\tau^2} + (\eta \Gamma_{jk}^2 + \xi^l \partial_l \eta \Gamma_{jk}^2) \left(\frac{dX^j}{d\tau} \frac{dX^k}{d\tau} + 2 \frac{dX^j}{d\tau} \frac{d\xi^k}{d\tau} \right) = 0.
\end{aligned}$$

In an inertial coordinate system, the inertial terms are zero. Therefore, the system of equations in the linear longitudinal dynamics in the ultra-relativistic regime is

$$\begin{aligned}
& \frac{d^2 \xi^0}{d\tau^2} + \frac{d\xi^2}{d\tau} \frac{dX^k}{d\tau} E_2 < y^m > \eta_{mk} + \frac{d\xi^k}{d\tau} \frac{dX^2}{d\tau} E_2 < y^m > \eta_{mk} + \\
& + 2\xi^l \partial_l \left(\frac{dX^k}{d\tau} \frac{dX^2}{d\tau} E_2 < y^m > \eta_{mk} \right) = 0 \\
& \frac{d^2 \xi^2}{d\tau^2} - \frac{d\xi^2}{d\tau} \frac{dX^k}{d\tau} E_2 < y^m > \eta_{mk} - \frac{d\xi^k}{d\tau} \frac{dX^0}{d\tau} E_2 < y^m > \eta_{mk} - \\
& - 2\xi^l \partial_l \left(\frac{dX^j}{d\tau} \frac{dX^0}{d\tau} E_2 < y^m > \eta_{mj} \right) = 0 \\
& \frac{d^2 X^1}{d\tau^2} = 0, \\
& \frac{d^2 X^2}{d\tau^2} = 0.
\end{aligned}$$

If $\epsilon^k = < y^k > - \frac{dX^k}{d\tau} \approx 0$ and since the distribution function has support on the unit hyperboloid, $< y^k > \frac{dX_k}{d\tau} \approx 1 + \alpha$. Using also the decoupling condition $\frac{dX^k}{d\tau} \frac{d\xi_k}{d\tau} \approx 0$, we have that

$$\frac{d^2 \xi^0}{d\tau^2} + \frac{d\xi^2}{ds} E_2 + 2\xi^l \partial_l \left(\frac{dX^2}{d\tau} E_2 \right) = 0, \quad (6.5.1)$$

$$\frac{d^2 \xi^2}{d\tau^2} - \frac{d\xi^0}{d\tau} E_2 - 2\xi^l \partial_l \left(\frac{dX^0}{d\tau} E_2 \right) = 0, \quad (6.5.2)$$

$$\frac{d^2 X^1}{d\tau^2} = 0, \quad (6.5.3)$$

$$\frac{d^2 X^2}{d\tau^2} = 0. \quad (6.5.4)$$

The only non-trivial equation has the form

$$\frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau} E_2 - 2\xi^l \partial_l \left(\frac{dX^0}{d\tau} E_2 \right) = 0.$$

In the ultra-relativistic limit the velocity field $\frac{dX^0}{d\tau} = \gamma$ (in units where the speed of light is equal to 1). Therefore, the equation above can be written as

$$\frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau} E_2 - 2\gamma \xi^l \partial_l E_2 (X + \xi^2) = 0. \quad (6.5.5)$$

We perform the following approximation in equation (6.5.5):

$$E_2(X + \xi^2) = E_2(X) + \xi^k \frac{\partial}{\partial \xi^k} E_2.$$

Due to the translational invariance of the partial derivatives, $\partial_l \equiv \frac{\partial}{\partial \xi^l}$ in the above expressions, by the chain rule. Then we have

$$\frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau} E_2 - 2\gamma \xi^k \frac{\partial}{\partial \xi^k} (E_2(X + \xi) - E_2(X)) = 0.$$

If $E_2(X + \xi)$ can be approximated linearly on ξ , using Euler's theorem of homogeneous functions one gets the following expression:

$$\frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau} E_2(\tau) - 2\gamma(\tau) (E_2((X + \xi)) - E_2(X)) = 0.$$

6.5.1 Examples

1. Constant longitudinal electric field.

In this case the equation of motion is

$$\frac{d^2 \xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau} E_2 = 0.$$

A particular solution is

$$\xi^2 = -\frac{\xi^2}{E_2}(e^{-E_2(\tau-\tau_0)} - 1).$$

2. **Alternate longitudinal electric field.** In this case, the electric field is

$$E_2(X^2 + \xi^2) = E_2(0)\sin(w_{rf}(X^2 + \xi^2)).$$

The differential equation is

$$\frac{d^2\xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau}E_2(0)\sin(w_{rf}(X^2 + \xi^2)) - 2\gamma E_2(0)(\sin(w_{rf}(X^2 + \xi^2)) - \sin(w_{rf}X^2)) = 0.$$

We can expand this equation in ξ , since ξ is small

$$\frac{d^2\xi^2}{d\tau^2} + \frac{d\xi^2}{ds}E_2(0)(\sin(w_{rf}X^2) + \cos(w_{rf}X^2)\xi^2) - 2\gamma E_2(0)(\cos(w_{rf}X^2)\xi^2) = 0.$$

At first order in ξ^2 we have the equivalent expression

$$\frac{d^2\xi^2}{d\tau^2} + \frac{d\xi^2}{d\tau}E_2(0)\sin(w_{rf}X^2) - 2\gamma E_2(0)(\cos(w_{rf}X^2)\xi^2) = 0.$$

We choose the initial phase such that $\sin(w_{rf}X^2) \simeq 0$; therefore $\cos(w_{rf}X^2) \simeq 1$ and the equation is

$$\frac{d^2\xi^2}{d\tau^2} - 2\gamma E_2(0)\xi^2 = 0. \tag{6.5.6}$$

This equation is a linearized version of the ordinary linearized longitudinal dynamics and has exactly the same structure [11].

6.6 Conclusion

We have seen that the linear beam dynamics in accelerator physics can be obtained from the Jacobi equation of the averaged Lorentz connection for electromagnetic fields $\mathbf{F}(\xi)$ linear in ξ . In particular we have proved that in the case where the

magnetic fields are linear on the deviation ξ , like in a dipole and quadrupole magnetic fields, the transverse dynamics can be interpreted as the dynamics of the Jacobi equation. A similar conclusion follows for the so-called longitudinal dynamics.

The advantages of this derivation are that it involves only observable quantities. Another advantage is that the averaged dynamics is linked through the distribution function to the collective behavior of the system.

The *off-set* effect calculated in *section 6.4* can be of relevance in the control and diagnostics of the beam parameters, since it is a direct observable quantity and because it is related with the collective description of the bunch of particles.

Chapter 7

Conclusions

7.1 General conclusions

This thesis describes the foundations of the averaged Lorentz force equation and its applications in the mathematical modeling of ultra-relativistic bunches of charged particles. Through the averaged Lorentz dynamics, a new theoretical justification of the use of fluid models in beam dynamics has been obtained. We have seen that for relativistic dynamics and for narrow probability distribution functions (in the sense that the diameter α of the distribution function obtained using the Euclidean metric in the laboratory frame is very small compared with the mass of the particles at rest), it is justified to substitute the original kinetic model based on the Vlasov equation by an averaged charged cold fluid model. One can control this approximation in terms of the *energy* of the bunch, the diameter of the distribution α and the time of the evolution of the bunch, all these variables measured in the laboratory frame.

Our method does not provide a system of differential equations for the fluid models. Instead it provides estimates of the differential operations which occur in the definition of the charged cold fluid model. Given a particular physical situation one can decide whether the given model is satisfactory or is a bad approximation using those bounds.

The averaged model has been an essential tool in obtaining those results. The

reason is that the averaged Lorentz force equation is simpler than the Lorentz force equation. The existence of normal coordinate systems associated to the averaged connection $\langle {}^L\nabla \rangle$ has been crucial for the calculations performed in *chapter 5*.

There are some advantages using the averaged Lorentz equation in the description of the dynamics of a bunch of particles instead of the Lorentz force equation:

1. At the classical level the electromagnetic field is measured by the effect on charged point particles. In case of the electromagnetic interaction of charged particles with an external electromagnetic field, the Lorentz force equation is the geodesic equation of a complicated non-linear *almost-connection* (for the notion of *almost-connection* see the appendix). The equation can be simplified by considering the associated averaged Lorentz connection. The averaged Lorentz connection is simpler than the original one, since the averaged connection is an affine connection on \mathbf{M} . This property allows us to have important technical tools (in particular normal coordinates).
2. Since the averaged Lorentz connection is simpler than the original Lorentz connection, one can use it to perform numerical simulations of the dynamics of a bunch of particles. The simplified model must allow a better numerical implementation in the simulation of the dynamics of a bunch containing a large number of charged particles.

In a similar way, there are advantages using the averaged Vlasov equation instead of the original Vlasov equation:

1. The calculations using the averaged Vlasov equation can be simplified using normal coordinates systems. This is basically because the underlying averaged Lorentz connection is an affine and torsion free connection.
2. It involves only the low moments of the distribution function \tilde{f} . This dependence on low moments, together with the fact that \tilde{f} is an approximation of f in the regime when the dynamics is ultra-relativistic and the distributions are narrow, can explain why current fluid models need only to consider differential

equations for the first, second and third moments, while the fourth moments are neglected. That is, the justification for the truncation schemes in fluid models comes from the structure of the averaged Lorentz connection.

Since the metric approach to geometrization contains intrinsic problems, we have considered an alternative geometric treatment of the Lorentz force. The framework has been the theory of non-linear connections associated with second order differential equations. We have applied this theory to the Lorentz force equation (4.1.1).

However, starting a dynamical model from a differential equation can be insufficient for some purposes. For instance, one could not recover a canonical Hamiltonian formalism. From a geometric point of view, if one has only a differential equation, one can not speak of variational problems.

The Lorentz force equation has been interpreted as the non-linear Berwald connection of a spray vector field $L\chi$. However, the above points imply that one has to look for a variational formulation. A consistent formulation has been described in *section 4.3*. This formulation is technically complicated and requires sheaves and pre-sheaves theory. It is a non-metric approach to semi-Randers spaces.

7.1.1 Brief discussion of the main problem

The main problem formulated in *section 2.3* has been addressed: we have obtained a recipe such that we can decide when the charged cold fluid model is a good approximation or not to the underlying kinetic model, in the context of accelerator physics. The conclusion is that for the actual accelerator machines, the charged cold fluid model is a good approximation to the underlying Vlasov model and that it can be used in the description of the beam dynamics.

As a byproduct we have obtained an averaged Lorentz force equation which is simpler than the original Lorentz force equation. We have proved that in physical situations, the Lorentz force equation can be substituted by the averaged Lorentz equation.

7.2 Generalizations and open problems

Some open problems that the present thesis leaves for future investigation are the following:

1. The geometric averaged method can be applied to other dynamical systems. As an example, let us consider the structure of the Lorentz connection ${}^L\nabla$. If $\eta(y, y) = 1$, the connection coefficients are polynomial in y up to third order. Therefore, it could be interesting to consider it in a similar way as an effective connection that has semi-spray coefficients of the form:

$$G^i(x, y) = a^i{}_{jk}(x)y^jy^k + a^i{}_{jkl}(x)y^jy^ky^l + a^i{}_{jklm}(x)y^jy^ky^ly^m + a^i{}_{jklmn}(x)y^jy^ky^ly^my^n. \quad (7.2.1)$$

This is the simplest generalization of the semi-spray coefficients of the connection ${}^L\nabla$.

There are some restrictions on these semi-sprays coefficients

- (a) The resulting geodesic equations must be gauge invariant. This means that there is an intrinsic gauge symmetry transformation and the tensors depend only of gauge invariant quantities.
- (b) The resulting equation must be Lorentz invariant.
- (c) The study of the basic dynamics of these connections compatible with the laws of the Electrodynamics (in particular with the Larmor law [3, pg 469]). A second order dynamics of a charged point particle must be a particular class of these dynamics, since it is well known and experimentally checked.
- (d) It is well known that if the back-reaction force is taken into account, then the Lorentz-Dirac equation follows from a balance equation [1-4]. One can investigate if a generalization of the type (7.2.1) can accommodate a second order equation which remains second order, after considering back-reaction.

- (e) It could be interesting to clarify whose semi-sprays of the type (7.2.1) or possible generalizations of the Lorentz force coming from non-linear electrodynamics [53].
- 2. The method used in this thesis to justify the use of the charged cold fluid model is applicable to other fluid models, like the warm fluid model. Indeed one can discriminate which model is better in some particular application, depending on the diameter of the given distribution function and the energy of the beam.
- 3. One can use the averaged Lorentz equation as a model in numerical simulations. Since the structure of the averaged equation is simpler than the original equation, it could be convenient to use it in numerical simulations in beam dynamics.
- 4. We have assumed some technical hypotheses. Although these assumptions are well defined and hold for the physical examples that we have in mind, it could be interesting to reduce the number of assumptions, obtaining more general results.

Apart from these points, directed to the core of this thesis, there are several points which could deserve more attention:

- 1. The notion of *almost-connection*. As is explained in the appendix, it is a natural generalization of the notion of connection. We think that it is non-trivial, since the non-extensibility of the covariant derivatives and parallel transport is a difficult property to prove.
- 2. The notion of semi-Randers space [14, work in progress]. This a basic notion that we have need to discuss but which is of interest on its own. Also it is interesting to generalize the notion of semi-Rander space associated with non-abelian symmetries.
- 3. The notions of structural stability introduced in *section 6.4*. It indicates a topological structure behind current fluid models which deserves investigation.

Appendix A

Mathematical appendix

A.1 Proofs for Chapter 3

The following proofs are adapted from reference [22].

Proof of proposition 3.3.2. The consistency of the equation (3.3.4) is proved in the following way,

$$\begin{aligned} \langle \nabla \rangle_X f &= \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* f \rangle_u = \\ &= \langle \pi_2|_u h_u(X) (\pi_v^* f) \rangle_u = \langle \pi_2|_u \pi_u^* (X(f)) \rangle_u = \langle Xf \rangle_u = Xf. \end{aligned}$$

The fourth equality holds because the definition of the horizontal local basis $\{\frac{\delta}{\delta x^0}|_u, \dots, \frac{\delta}{\delta x^{n-1}}|_u\}$ in terms of $\{\frac{\partial}{\partial x^i}, i = 0, \dots, n-1\}$ and $\{\frac{\partial}{\partial y^i}, i = 0, \dots, n-1\}$.

We check the properties characterizing a linear covariant derivative associated with the averaged connection $\langle \nabla \rangle$:

1. $\langle \nabla \rangle_X$ is a linear application acting on sections of **TM**:

$$\langle \nabla \rangle_X (Y_1 + Y_2) = \langle \nabla \rangle_X Y_1 + \langle \nabla \rangle_X Y_2$$

$$\langle \nabla \rangle_X \lambda Y = \lambda \tilde{\nabla}_X Y,$$

$$\forall Y_1, Y_2, Y \in \Gamma \mathbf{TM}, \lambda \in \mathbf{R}, \quad X \in \mathbf{T}_x \mathbf{M}. \quad (\text{A.1.1})$$

For the first equation, the proof consists in the following calculation,

$$\begin{aligned} \langle \nabla \rangle_X (Y_1 + Y_2) &= \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* (Y_1 + Y_2) \rangle_v = \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* Y_1 \rangle_v + \\ &+ \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^* Y_2 \rangle_v = \langle \nabla \rangle_X Y_1 + \langle \nabla \rangle_X Y_2. \end{aligned}$$

For the second condition we have that

$$\langle \nabla \rangle_X (\lambda Y) = \langle \pi_2|_u \nabla_{\iota_u(X)} \pi_v^* (\lambda Y) \rangle_v = \lambda \langle \pi_2|_u \nabla_{\iota_u(X)} \pi_v^* Y \rangle_v = \lambda \langle \nabla \rangle_X Y.$$

2. $\langle \nabla \rangle_X Y$ is a \mathcal{F} -linear with respect to X :

$$\langle \nabla \rangle_{X_1+X_2} Y = \langle \nabla \rangle_{X_1} Y + \langle \nabla \rangle_{X_2} Y,$$

$$\langle \nabla \rangle_{fX} (Y) = f(x) \langle \nabla \rangle_X Y,$$

$$\forall Y \in \mathbf{TM}, v \in \pi^{-1}(z), \quad X, X_1, X_2 \in \mathbf{T}_x \mathbf{M}, f \in \mathcal{F}(\mathbf{M}). \quad (\text{A.1.2})$$

To prove the first equation is enough the following calculation:

$$\begin{aligned} \langle \nabla \rangle_{X_1+X_2} Y &= \langle \pi_2|_u (\nabla_{h_u(X_1+X_2)}) \pi_v^* Y \rangle_v = \\ &= \langle \pi_2|_u \nabla_{h_u(X_1)} \pi_v^* Y \rangle_v + \langle \pi_2|_u \nabla_{h_u(X_2)} \pi_v^* Y \rangle_v = \\ &= (\langle \nabla \rangle_{X_1} Y) + (\langle \nabla \rangle_{X_2} Y). \end{aligned}$$

For the second condition the proof is similar.

3. The Leibnitz rule holds:

$$\langle \nabla \rangle_X (fY) = df(X)Y + f \langle \nabla \rangle_X Y, \quad \forall Y \in \Gamma \mathbf{TM}, f \in \mathcal{F}(\mathbf{M}), \quad X \in \mathbf{T}_x \mathbf{M}, \quad (\text{A.1.3})$$

where $df(X)$ is the action of the 1-form $df \in \Lambda^1 \mathbf{M}$ on $X \in \mathbf{T}_x \mathbf{M}$ In order to

prove (3.3.7) we use the following property:

$$\pi_v^*(fY) = \pi_v^*f\pi_v^*Y, \quad \forall Y \in \Gamma\mathbf{TM}, f \in \mathcal{F}(\mathbf{M}).$$

Then

$$\begin{aligned} \langle \nabla \rangle_X (fY) &= \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^*(fY) \rangle_u = \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^*(f) \pi_v^*Y \rangle_u = \\ &= \langle \pi_2|_u (\nabla_{h_u(X)} (\pi_v^*f)) \pi_v^*(Y) \rangle_u + \langle \pi_2|_u (\pi_v^*f) \nabla_{h_u(X)} \pi_v^*(Y) \rangle_u = \\ &= \langle \pi_2|_u (h_u(X)(\pi_v^*f)) \pi_v^*(Y) \rangle_u + f_x \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^*(Y) \rangle_u = \\ &= \langle (X_x f) \pi_2|_u \pi_u^*(Y) \rangle_u + f_x \langle \pi_2|_u \nabla_{h_u(X)} \pi_v^*(Y) \rangle_u. \end{aligned}$$

For the first term we perform the following simplification,

$$\begin{aligned} \langle (Xf) \pi_2|_u \pi_u^*(Y) \rangle_u &= (Xf) \langle \pi_2|_u \pi_u^*(Y) \rangle_u = \\ &= (Xf) (\langle \pi_2|_u \pi_u^* \rangle_u) Y = (Xf) (\langle I \rangle_u) Y. \end{aligned}$$

Returning to the above calculation, we obtain

$$\tilde{\nabla}_X(fY) = \tilde{\nabla}_X(f)Y + f\tilde{\nabla}_X Y = df(X)Y + f\tilde{\nabla}_X Y.$$

□

The generalization of $\langle {}^L\nabla \rangle$ to higher order tensor bundles is as usual.

Proof for Proposition 3.3.4 and corollary 3.3.5

$$\begin{aligned} T_{\langle \nabla \rangle}(X, Y) &= \langle \pi_2|_u \nabla_{h_u(X)} \pi_w^* \rangle_u Y - \langle \pi_2|_u \nabla_{h_u(Y)} \pi_w^* \rangle_u X - [X, Y] = \\ &= \langle \pi_2|_u \nabla_{h_u(X)} \pi_u^* \rangle_u Y - \langle \pi_2|_u \nabla_{h_u(Y)} \pi_u^* \rangle_u X \\ &\quad - \langle \pi_2|_u \pi_u^*[X, Y] \rangle_u = \end{aligned}$$

$$=< \pi_2|_u (\nabla_{h_u(X)} \pi^* Y - \nabla_{h_u(Y)} \pi^* X - \pi^*[X, Y]) >_u = < T(X, Y) > .$$

On the other hand:

$$T_{<\nabla>}(X, Y) = < \pi_2|_u \nabla_{h_u(X)} \pi_w^* >_u Y - < \pi_2|_u \nabla_{h_u(Y)} \pi_w^* >_u X - [X, Y] =$$

$$=< \pi_2|_u \nabla_{h_u(X)} \pi_u^* >_u Y - < \pi_2|_u \nabla_{h_u(Y)} \pi_u^* >_u X - < \pi_2|_u \pi_u^*[X, Y] > =$$

$$=< \pi_2|_u (\nabla_{h_u(X)} \pi^* Y - \nabla_{h_u(Y)} \pi^* X - \pi^*[X, Y]) >_u = 0. \square$$

A.2 Ordinary differential equations

We state here a result on differential equations that we have used several times in the text. It was used in [21] to show the local existence and uniqueness of parameterized geodesics and similar results on the existence and uniqueness of solutions of ordinary differential equations,

Theorem A.2.1 *Let $f_i(t, y, s)$ be a family of n functions defined in $|t| < \delta$ and $(y, s) \in \mathbf{D}$, where \mathbf{D} is an open set in \mathbf{R}^{n+m} . If $f_i(t, y, s)$ are continuous in t and differentiable of class \mathcal{C}^1 in y , then there exists a unique family $\phi(t, y, s)$ of n functions defined in $|t| < \delta$ and $(\eta, s) \in \mathcal{D}$, where $0 < \delta < \delta$ and \mathcal{D} is an open subset of \mathbf{D} such that*

1. $\phi(t, \eta, s)$ is differentiable of class \mathcal{C}^1 in t and η .
2. $\frac{\partial \phi_i}{\partial t} = f_i(t, \phi(t, \eta, s), s)$.
3. $\phi(0, \eta, s) = \eta$.

If $f(t, y, s)$ is differentiable of class \mathcal{C}^p , $0 \leq p \leq \infty$, in y and s , then $\phi(t, \eta, s)$ is differentiable of class \mathcal{C}^{p+1} in t and of class \mathcal{C}^q in η and s .

The main use in the thesis was the following. Assume that the geodesic equations are of the form

$$\frac{dy^i}{dt} = G^i(x, y, s); \quad \frac{dx^i}{dt} = y^i, \quad i = 1, \dots, n.$$

We have applied the theorem to the case where s defines a smooth homotopy. Therefore, we have to apply to the case where s is 1-dimensional and $f = (y^i, G^i)$, with $i = 1, \dots, 2n$.

We have used this theorem to prove smoothness properties of the solutions of several differential equations.

A.3 Basic notions of asymptotic analysis

In this *appendix* we follow the notation of [43, *chapter 1*].

Definition A.3.1 For two complex functions $f(z)$ and $g(z)$ we write $f(z) = \mathcal{O}(g(z))$ as $z \rightarrow z_0$ iff there is a positive constants K and c such that for $0 < |z - z_0| < c$ one has that $|f| \leq K|g|$.

Definition A.3.2 An infinite sequence of functions $\{\Phi_n(z), n = 1, 2, \dots\}$ is an asymptotic sequence as $z \rightarrow z_0$ if

$$\lim_{z \rightarrow z_0} \frac{\Phi_{n+1}(z)}{\Phi_n(z)} = 0, \quad \forall n \in \mathbf{N}.$$

Definition A.3.3 Given the asymptotic sequence $\{\Phi_n(z), n = 1, 2, \dots\}$ as $z \rightarrow z_0$ the expression

$$f(z) = \sum_{n=1}^{N-1} a_n \Phi_n(z) + \mathcal{O}(\Phi_N(z)) \text{ as } z \rightarrow z_0,$$

is said to be an asymptotic expansion of the function $f(z)$ in terms of $\{\Phi_n(z), n = 1, 2, \dots\}$.

Given an asymptotic sequence $\{\Phi_n(z), n = 1, 2, \dots\}$, if for a given function f the asymptotic expansion exists it is unique, with coefficients given by

$$a_k = \lim_{z \rightarrow z_0} \left\{ \frac{f(z) - \sum_{n=1}^{k-1} a_n \Phi_n(z)}{\Phi_k} \right\}.$$

Asymptotic expansions have the following elementary properties:

1. The first order term in the expansion is the leading term, which provides the major contribution to the series.
2. The asymptotic expansion of a function depends on the choice of the asymptotic sequence.
3. The asymptotic expansion as $z \rightarrow z_0$ is a linear operation respect to the function which is being expanded: $f \mapsto (a_1, \dots, a_{N-1})$ is linear in f , $\forall N > 1$.

4. If the derivative of the function f has an asymptotic expansion, the asymptotic expansion of the derivative of a function is the derivative term by term of the expansion of f .
5. The asymptotic expansion of the real integral of a function is the real integral of the expansion, integrated term by term.
6. The product of asymptotic expansions is in general non-asymptotic. However, the product of asymptotic expansions in power series around the same point is also an asymptotic expansion.

If the asymptotic expansion is convergent, there is a convergence region such that the approximation of a function by an asymptotic series is becoming more accurate when we consider more terms in the expansion. If the asymptotic series is a power series and divergent, the better accuracy attainable by an expansion consists on taking the expansion

$$f(z) \sim \sum_{n=1}^N (z)_{n=1} a_n \Phi_n(z)$$

such that $N(z)$ is the last term that the magnitude of the terms $|a_n \Phi_n|$ is decreasing with n . After this term $a_N \Phi_N$, the rest of the terms are increasing. Usually, the error in the expansion in a power series is of the order of the first term neglected.

A.4 Notion of almost-connection

During our analysis of the Lorentz force equation, we have considered the associated non-linear connection on \mathbf{TN} and the associated linear connections on $\pi^*\mathbf{TM} \rightarrow \Sigma$. However, strictly speaking the system of differential equations (4.1.1) does not define a non-linear connection on $\mathbf{T}\Sigma$. The reason is the following. Let us fix a point $u \in \Sigma$. There exists a natural embedding $e : \Sigma \hookrightarrow \mathbf{TM}$. One can consider the sub-bundle $e(\Sigma) \hookrightarrow \mathbf{TM}$ and the *covariant derivative* ${}^L D$ acting on elements of $\Gamma \mathbf{T}e(\Sigma)$. This covariant derivative can be extended to derive sections of the extended bundle $\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{N}$, which is a sub-bundle of \mathbf{TM} :

$${}^L \hat{D} : \Gamma(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{N}) \times \Gamma(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{N}) \longrightarrow \Gamma(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{N}),$$

$${}^L \hat{D}_{\hat{X}} \hat{Y} = \hat{X}^i \frac{\partial \hat{Y}^j}{\partial x^i} \frac{\partial}{\partial x^j} + {}^L \Gamma^i_{jk} \hat{Y}^k \hat{X}^j \frac{\partial}{\partial x^i}, \quad \hat{X}, \hat{Y} \in \Gamma(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{TM}).$$

From the definition of ${}^L D$, this operator cannot be extended in a smooth and natural way to be a covariant derivative on \mathbf{TM} . This is because of the appearance of factors $\sqrt{\eta(y, y)}$ in the connection coefficients: the function $\sqrt{\eta(y, y)}$ is not defined in the whole \mathbf{TM} . The non-extendibility of this function to \mathbf{TM} is the origin of the problem to extend the covariant derivative along arbitrary directions in \mathbf{TM} .

The above fact suggests the existence of a mathematical object which is a generalization of the ordinary notion of covariant derivative in the sense that allows covariant derivatives along *outer directions to the manifold*, although not obtained as a restriction of an *ambient covariant derivative* operator.

Example. The Lorentz connection was obtained from the Lorentz force equation assuming that the torsion is zero. We obtained the following connection coefficients:

$${}^L \Gamma^i_{jk} = \eta \Gamma^i_{jk} + T^i_{jk} + L^i_{jk},$$

$$L^i_{jk} = \frac{1}{2\eta(y, y)} (\mathbf{F}^i_{jy} y^m \eta_{mk} + \mathbf{F}^i_{ky} y^m \eta_{mj}),$$

$$T^i{}_{jk} = \mathbf{F}^i{}_m \frac{y^m}{\sqrt{\eta(y, y)}} \left(\eta_{jk} - \frac{1}{\eta(y, y)} \eta_{js} \eta_{kl} y^s y^l \right).$$

This defines a rule to derive sections of $\mathbf{T}\Sigma$ along directions of $\mathbf{T}\Sigma$. The same coefficients provide a rule to derive sections of $\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{TM}$ along directions of \mathbf{TM} , but we can extend the definition of the covariant derivative acting on sections of \mathbf{TTM} .

A related issue is the following. On the unit hyperboloid recall that $T^i{}_{jk} y^j y^k = 0$. Therefore, if one considers instead an alternative connection $\tilde{\nabla}$ defined by the connection coefficients

$${}^L \tilde{\Gamma}^i{}_{jk} = \eta \Gamma^i{}_{jk} + L^i{}_{jk},$$

the corresponding geodesic equation (parameterized by the proper time of the Lorentzian metric η) is again the Lorentz force equation [49]. Therefore we see that both ${}^L \nabla$ and $\tilde{\nabla}$ reproduce the Lorentz force and both are torsion-free (the connection coefficients are symmetric in the lower indices). This is in contradiction with the fact that a connection of Berwald type (linear or non-linear) is determined by the set of all geodesics as parameterized curves on the base manifold \mathbf{M} and the torsion tensor (for linear connections the procedure can be seen in [35]; for non-linear connections, a procedure to define the connection is described for example in [34]).

A solution to this dilemma comes from the fact that the Lorentz force equation applies only to time-like trajectories for which the tangent velocity vector fields live on the unit hyperboloid Σ ; that is, we cannot extend the operator

$${}^L \hat{D} : \Gamma\left(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{TM}\right) \times \Gamma\left(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{TM}\right) \longrightarrow \Gamma\left(\bigsqcup_{u \in \Sigma} \mathbf{T}_u \mathbf{TM}\right)$$

to a genuine operator of the form

$${}^L D : \Gamma(\mathbf{TM}) \times \Gamma(\mathbf{TM}) \longrightarrow \Gamma(\mathbf{TM}).$$

In other words, we cannot extract enough information from the Lorentz force equation to determine a *projective connection* [44], because on the base manifold, we do not have information about all the possible geodesics. This is why we cannot

determine the connection.

One can be tempted to extend the Lorentz force equation to another equation valid for any kind of trajectory. This is partially accomplished by the averaged connection. However, these extensions could be non-natural or non unique and indeed hide a natural object such like *almost-connection*.

Preliminary Definition of Almost-Connection

Let \mathbf{M} be a manifold of dimension n . An almost-connection is a spray vector field $\chi \in \Gamma\mathbf{TM}$ defined on a sub-bundle $\mathbf{TD} \hookrightarrow \mathbf{TTM}$, with \mathbf{TD} a sub-bundle of arbitrary co-dimension. Associated with \mathbf{D} is the corresponding almost projective covariant derivative on $\pi^*\mathbf{TM}$.

Examples.

1. A projective connection (in the sense of Cartan [44]) is an almost connection such that $\mathbf{D} = \mathbf{TM}$.
2. The Lorentz connection ${}^L\nabla$ provides an example where $\Sigma = \mathbf{D} \neq \mathbf{TM}$. One can define a *Koszul connection* acting on $\Gamma(\bigsqcup_{u \in \Sigma} \mathbf{T}_u\mathbf{TM})$.
3. The averaged covariant derivative $\langle {}^L\nabla \rangle$ is an affine connection and therefore an almost connection in the above sense.

One can consider the corresponding linear connections ${}^L\nabla$ and $\tilde{\nabla}$ on $\pi^*\mathbf{TM}$. Generally, their averaged connections $\langle {}^L\nabla \rangle$ and $\langle \tilde{\nabla} \rangle$ (introduced in *section 4.7*) are not the same. However, we would like to have an averaged operation which is well defined for the objects in a given category. By definition this will be the category of *almost-connections* and the corresponding morphisms. We require that the result of the averaging operation be the same for each representative belonging to the same almost connection,

Definition A.4.1 *Let \mathbf{M} be a manifold of dimension n . A non-linear almost connection is the maximal set of semi-sprays χ defined on a sub-bundle $\mathbf{TD} \subset \mathbf{TTM}$ such that they have the same averaged linear covariant derivative $\langle \nabla \rangle$ and the*

same torsion tensor

$$\hat{T}(\hat{X}, \hat{Y}) := {}^L\hat{D}_{\hat{X}}\hat{Y} - {}^L\hat{D}_{\hat{Y}}\hat{X} - [\hat{X}, \hat{Y}], \quad \hat{X}, \hat{Y} \in \Gamma\mathbf{TD}. \quad (\text{A.4.1})$$

Associated with χ over \mathbf{D} is the corresponding *almost-covariant derivative* on the bundle $\pi^*\mathbf{TM} \longrightarrow \mathbf{D}$. Any of the connections in the same almost-connection has the same averaged connection. The distance function (4.6.3) can also being defined for almost-connections.

The general properties of *almost connections* are being explored in a separate work. Some of these properties are based on straightforward generalizations of the quantities associated with Koszul connections. For instance, the generalization of the curvature tensor is

$$\hat{R}(\hat{X}, \hat{Y}, \hat{Z}) := {}^L\hat{D}_{\hat{X}}{}^L\hat{D}_{\hat{Y}}\hat{Z} - {}^L\hat{D}_{\hat{Y}}{}^L\hat{D}_{\hat{X}}\hat{Z} - {}^L\hat{D}_{[\hat{X}, \hat{Y}]}\hat{Z}, \quad \hat{X}, \hat{Y}, \hat{Z} \in \Gamma\mathbf{TD}. \quad (\text{A.4.2})$$

It is not easy to handle a notion of parallel transport for almost-connections, since in general there will be initial conditions such that the auto-parallel curve goes out from the sub-bundle \mathbf{D} , even for arbitrary short-time parallel transports. Indeed, for auto-parallel curves whose initial velocity vector is not on \mathbf{TD} , the trajectory *goes out* from \mathbf{D} after any finite time. This point is related with the notion of general connection [62]. However, the notion of almost-connection is even more general, since the solution of the projections on \mathbf{N} condition $\nabla_{h(\xi)}\pi^*(\xi) = 0$ could not exists. For a generalized connection, this projection always exist.

A.5 Basic notions of Sobolev spaces

Sobolev spaces are complete normed vector spaces (therefore Banach spaces) where the norms measure also the derivatives of the function. We have used Sobolev norms in *chapter 5* to introduce the bounds on some differential expressions appearing in the averaged Vlasov model. In this appendix we provide the basic notions of Sobolev norms and some additional notions of Sobolev spaces to understand the meaning of these expressions and its implications for further generalizations. We will follow references [41] and [57] because of their clarity in exposition. We assume that the theory of Lebesgue's integral holds.

Let Ω be an open set of a manifold. There are some basic definitions:

Definition A.5.1 *Two functions $f, g : \Omega \longrightarrow \mathbf{R} \cup \{\pm\infty\}$ are equivalent iff they are equal almost everywhere on Ω , that is, the sub-set $A \subset \Omega$ where $g \neq f$ is a null set.*

Definition A.5.2 *Let Ω be open, $p \geq 1$ ($p \in \mathbf{R}$). $L^p(\Omega)$ is the set of all Lebesgue measurable functions $f : \Omega \longrightarrow \mathbf{R} \cup \{\pm\infty\}$ for which $|f|^p$ is integrable over Ω . For $f \in L^p(\Omega)$ we set*

$$\|f\|_{L^p(\Omega)} := \left(\int_{\Omega} |f|^p dx \right)^{\frac{1}{p}}.$$

In order to define Sobolev norms, we introduce weak derivatives.

Definition A.5.3 *Let $f \in L^1(\Omega)$. A function $v \in \mathbf{L}^1(\Omega)$ is called the weak derivative of f in the direction x^i if*

$$\int_{\Omega} v(x) \phi(x) dx = - \int_{\Omega} f(x) \frac{\partial \phi(x)}{\partial x^i} dx.$$

Definition A.5.4 *Let $f \in L^1(\Omega)$, $\beta := (\beta_1, \dots, \beta_d)$ with $\beta_i \geq 0$ ($i = 1, \dots, d$), $|\beta| := \sum_{i=1}^d \beta_i > 0$. Then*

$$D_{\beta} \phi := \left(\frac{\partial}{\partial x^1} \right)^{\beta_1} \cdots \left(\frac{\partial}{\partial x^d} \right)^{\beta_d} \phi, \quad f \in \mathcal{C}^{|\beta|}(\Omega).$$

Definition A.5.5 *A function $v \in L^1(\Omega)$ is called β -weak derivative of f and written*

$v = D_\beta f$ if

$$\int_{\Omega} v(x)\phi(x) dx = (-1)^{|\beta|} \int_{\Omega} f(x)D_\beta\phi(x) dx.$$

We can now define Sobolev spaces and Sobolev norms:

Definition A.5.6 For $k \in \mathbf{N}$ a natural number, $1 \leq p \leq \infty$, we define the Sobolev space $\mathcal{W}^{k,p}(\Omega)$ by $\mathcal{W}^{k,p}(\Omega) := \{f \in L^p(\Omega) \mid D_\beta f \text{ exists and is in } L^p(\Omega)\}$. The Sobolev norms are defined by

$$\|f\|_{k,p} := \left(\sum_{|\beta| \leq k} \int_{\Omega} |D_\beta f(x)| dx \right)^{\frac{1}{p}}, \quad (\text{A.5.1})$$

and by

$$\|f\|_{k,p} := \sum_{|\beta| \leq k} \max \sup_{x \in \Omega} |D_\beta f(x)|. \quad (\text{A.5.2})$$

A.5.1 Basic properties of Sobolev spaces

Some properties of Sobolev spaces are the following:

Theorem A.5.7 Let $k \in \mathbf{N}$ be a natural number and $1 \leq p < \infty$. Then the following is true

1. The normed space $(\mathcal{W}^{k,p}(\Omega), \|f\|_{k,p})$ is a Banach space.
2. Let $f \in \mathcal{W}^{k,p}(\Omega)$ and $\theta \in \mathcal{C}^1(\mathbf{R})$. Then Poincare's inequality holds:

$$|f|_{(0,p)} \leq C(\Omega, p) |f|_{1,p}, \quad f \in (\mathcal{W}_c^{1,p}(\Omega)), \quad (\text{A.5.3})$$

with $C(\Omega, p)$ a constant and $\mathcal{W}_c^{1,p}(\Omega)$ the completion of the space of smooth functions on Ω with compact support.

3. $\mathcal{W}_c^{p,p}(\Omega)$ is a Hilbert space.

There is a relevant result (Sobolev embedding theorem) which gives sufficient conditions to embed Sobolev spaces in L^q spaces, in spaces of continuous functions or in spaces with some regularity conditions [41,57,58]. We do not need this theorem

here, but it could be important for of further generalizations of the results of *chapter 5*.

A.5.2 Sobolev spaces of functions defined on manifolds

The discussion before of Sobolev spaces is restricted to open domains. One can extend the definition to manifolds. First one reduces to an open domain $\mathbf{U} \subset \mathbf{M}$. If the manifold is differentiable, there are partitions of the unity [32]. First, one can speak of $\mathcal{W}_c^{1,p}(\mathbf{U})$. Using an atlas of the manifold and a associated partition of the unity, we can define the Sobolev norms $(\mathcal{W}^{k,p}(\mathbf{M}), \|f\|_{k,p})$ from $(\mathcal{W}^{k,p}(\Omega), \|f\|_{k,p})$ in the usual way as the integrals are defined over manifolds from the local description using coordinates neighborhoods [32, *chapter 4*].

A.5.3 Examples of Sobolev spaces

There are some Sobolev norms that we have used in *chapter 5*. These are:

1. $(\mathcal{W}^{1,1}(\mathbf{M}), \|f\|_{1,1})$. This is a Hilbert space, with a norm defined by the function

$$\|f\|_{\mathcal{W}^{1,1}} := \left(\sum_{|\beta| \leq 1} \int_{\Omega} |D_{\beta} f(x)| dx \right) = \int_{\Omega} (|f(x)| + |\sum_i \partial_i f|) dx.$$

2. $(\mathcal{W}^{0,2}(\mathbf{M}), \|f\|_{0,2})$. The norm of a function is defined as

$$\|f\|_{0,2} := \left(\int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}.$$

Bibliography

- [1] P. A. M. Dirac, *Classical Theory of Radiating Electrons*, Proceedings of the Royal Society of London, vol: A137, 148-169 (1938).
- [2] L. D. Landau, E. M. Lifshitz, *The Classical Theory of Fields*, Pergamon, Oxford (1962).
- [3] J. D. Jackson, *Classical Electrodynamics*, third ed. Wiley (1998).
- [4] F. Rohrlich, *Classical Charged Particles*, Addison-Wesley, Redwood City (1990).
- [5] E. Poisson, *An introduction to the Lorentz-Dirac equation*, arXiv:gr-qc/9912045.
- [6] D. A. Burton, J. Gratus, R. W. Tucker, *Asymptotic analysis of ultra-relativistic charge*, Annals of Physics 322 (2007) 599-630.
- [7] J. Ehlers, *General Relativity and Kinetic Theory*, Proceedings of the International Summer School of Physics Enrico Fermi, pg 1-70 (1971).
- [8] R. C. Davidson, *Physics of Nonneutral Plasmas*, World Scientific (2001).
- [9] H. Wiedemann, *Particle Accelerator Physics I*, Second Edition, Springer (1998).
- [10] B. Holzer, *Introduction to Transverse Beams Optics*, Post-Graduate Lectures at the Cockcroft Institute, September (2007).
- [11] J. Le Duff, *Longitudinal Dynamics*, Post-Graduate Lectures at the Cockcroft Institute (2007).

- [12] D. J. Saunders, *A new approach to the nonlinear connection associated with a second-order (and higher-order) differential equation fields*, J. Phys. A: Math. Gen. **30** (1997) 1739-1743.
- [13] M. Crampin, *On Horizontal Distributions on the Tangent Bundle of a Differentiable manifold*, J. London Math. Soc. (2),3 (1971), 178-182.
- [14] R. Gallego Torrome, *On the Notion of Semi-Randers Spaces*, arXiv:0906.1940.
- [15] R. Gallego Torrome, *Geometric Formulation of the Classical Dynamics of Charged Particles in a External Electromagnetic Field*, arXiv:0905.2060.
- [16] R. Gallego Torrome, *Fluid Models from Kinetic Theory using Geometric Averaging*, arXiv:0912.2767; *Averaged Lorentz Dynamics and an application in Plasma Dynamics*, arXiv:0912.0183.
- [17] I.M. Benn, R.W. Tucker, *An introduction to Spinors and Geometry with applications in Physics*, Bristol ; Philadelphia : A. Hilger (1987).
- [18] Jonathan Gratus, *Distributional solutions of the Vlasov's equation*, arXiv:0809.2944.
- [19] Robin W. Tucker, *Classical field-particle dynamics in space-time geometries*, Proc. R. Soc. Lond. A (2004) **460**, 2819-2844.
- [20] W. Greub, S. Halperin, R. Vanstone, *Connections, Curvature, and Cohomology, Vol I*, Academic Press (1972).
- [21] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, Vol I*, Wiley Interscience, New York (1969).
- [22] R. Gallego Torrome, *Averaged Structures Associated to a Finsler structure*, math.DG/0501058.
- [23] John K. Beem, P. E. Ehrlich, K. L. Easley, *Global Lorentzian Geometry*, Second Edition, CRC Press (1996).

- [24] D. Bao, S.S. Chern and Z. Shen, *An Introduction to Riemann-Finsler Geometry*, Graduate Texts in Mathematics 200, Springer-Verlag.
- [25] W. A. Poor, *Differential Geometric Structures*, MacGraw-Hill Book Company (1981).
- [26] G. Randers, *On an Asymmetrical Metric in the Four-Space of General Relativity*, Phys. Rev. **59** (1941), 195-199.
- [27] G. S. Asanov, *Finsler Geometry, Relativity and Gauge Theories*, D. Reidel, Dordrecht. (1985).
- [28] J. K. Beem, *Indefinite Finsler Spaces and Timelike Spaces*, Canad. J. Math. **22**, (1970) 1035.
- [29] J. Beem, *Characterizing Finsler Spaces which are Pseudo-Riemannian of Constant Curvature*, Pac. J. of Mathematics, vol. **64**, num. 1 (1976), 67.
- [30] V. Perlick, *Fermat Principle in Finsler Spacetimes*, Gen.Rel.Grav. **38** (2006) 365-380.
- [31] J. Douglas *Solution of the inverse problem of the calculus of variations*, Proceedings of the National Academy of Sciences 25, (1939) 631-637.
- [32] F. Warner *Foundations of Differentiable Manifolds and Lie Groups*, Scott, Foresman and Company (1971).
- [33] R. Hartshorne, *Algebraic Geometry*, Springer Verlag (1977).
- [34] R. Miron, D. Hrimiuc, H. Shimada, V. Sabau *The Geometry of Hamilton and Lagrange Spaces*, Fundamental Theories in Physics 118, Kluwer (2001).
- [35] N. J. Hicks, *Notes on Differential Geometry*, D. Van Nostrand Company (1965).
- [36] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of the Space-Time*, Cambridge Monographs on Mathematical Physics (1973).
- [37] P. Amendt, *Plasma wave equations of state*, Phys. Fluids **29** (5) (1986), 1458.

- [38] S. Pennisi and A. M. Anile, *Fluid Models for relativistic Electron Beam: an independent derivation*, Phys. Fluids B3(4) (1991), 1091.
- [39] P. Amendt, H. Weitzner, *Relativistic Covariant Warm Charge Fluid Model*, Phys. Fluids **28** (1985), 949.
- [40] C. Chicone, *Ordinary Differential Equations with Applications*, 2nd. Edition, Springer-Verlag (2006).
- [41] L. P. Lebedev, I. I. Vorovich and G. M. I Gladwell, *Functional Analysis*, 2nd. Edition , Kluwer Academic Publishers (2002).
- [42] W. Rudin, *Real and Complex Analysis*, Mac Graw-Hill, (1966).
- [43] J. D. Murray, *Asymptotic Analysis*, Applied Mathematical Siences, n. 48, Springer-Verlag (1973).
- [44] E. Cartan, *Sur les varietes a connexion projective*, Bulletin de la Société Mathématique 52 (1924), 205241 .
- [45] D. Bao, *On two curvature-driven problems in Riemann-Finsler geometry*, Advanced Studies in Pure Mathematics 48, 2007 Finsler Geometry, Sapporo 2005 - In Memory of Makoto Matsumoto pp. 1971.
- [46] M. Crampin, E. Martinez, *Towards a geometric understanding of Douglas's solution of the inverse problem of the calculus of variations*, Inverse Problems **10** (1994), 245-260.
- [47] B. O'Neill, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press (1983).
- [48] M. Berge, *A Panoramic view of Riemannian Geometry*, Springer Verlag (2003).
- [49] R. Miron, *Finsler-Lagrange Spaces with (α, β) -Metrics and Ingarden Spaces*, Reports on Mathematical Physics, Vol. **58** (2006), 417.
- [50] R. Gallego Torrome, *Geometric Formulation of the Averaging Operation*, arXiv:0905.3181.

- [51] N. Steenrod, *The Topology of Fibre Bundles*, Princeton University Press (1951).
- [52] A. Kolmogorov, S.V. Fomin, *Elements of the Theory of Functions and Functional Analysis*, Dover (1957).
- [53] M. Born, L. Infeld, *Foundations of a new Field Theory*, Proc. Roy. Soc. Lond. **A144** (1934), 425-451.
- [54] John Milnor, *Morse Theory*, Annals of Mathematical Studies, Princeton University Press (1969).
- [55] A. Magnus, S. Winkler, *Hill's equation*, Dover (1966).
- [56] M. I. Katsnelson, *Graphene: carbon in two dimensions*, arXiv: cond-mat/0612534.
- [57] J. Jost, *Postmodern Analysis*, Universitext Springer (1998).
- [58] R. A. Adams, *Sobolev Spaces*, Academic Press (1975).
- [59] T. D. Hodgkinson, *A modified equation of geodesic deviation*. Gen. Relativ. Gravit. 3 (1972), 351375.
- [60] R. K. Sachs, H. Wu, *General Relativity for Mathematicians*, Springer-Verlag (1977).
- [61] V. I. Arnold, *Geometric Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag (1983).
- [62] I. Kolár, P. W. Michor, J. Slovák, *Natural Operations in Differential Geometry*, Springer-Verlag (1993).